Sewer asset management : Impact of data quality and models' parameters on condition assessment of assets and asset stocks

Mehdi Ahmadi

N° d'ordre 2014ISAL0034 Année 2014

Thèse

GESTION PATRIMONIALE DES RESEAUX D'ASSAINISSEMENT: IMPACT DE LA QUALITE DES DONNEES ET DU PARAMETRAGE DU MODELE UTILISE SUR L'EVALUATION DE L'ETAT DES TRONÇONS ET DES PATRIMOINES

SEWER ASSET MANAGEMENT: IMPACT OF DATA QUALITY AND MODELS' PARAMETERS ON CONDITION ASSESSMENT OF ASSETS AND ASSET STOCKS

Présentée devant
L'institut national des sciences appliquées de Lyon

Pour obtenir
Le grade de docteur

Ecole doctorale : Mécanique, Energétique, Génie Civil, Acoustique (MEGA)
Spécialité : Génie Civil Environnemental

Par
Mehdi AHMADI
Ingénieur en « Génie civil», Université de Téhéran, Iran.
Ingénieur en « Génie Environnement », Ecole Nationale des Ponts et Chaussées, France.
Master en « Génie Civil et Environnement », Ecole Nationale des Ponts et Chaussées, France.

Soutenue le vendredi 11 Avril 2014 devant la Commission d'examen

Jury

Mme. BARRAUD Sylvie	Professeure (INSA de Lyon)	Examinatrice
M. BREUL Pierre	Professeur (Univ. Blaise Pascal, Clermont-Ferrand)	Examinateur
M. BREYSSE Denys	Professeur (Université Bordeaux 1)	Examinateur
M. CHERQUI Frédéric	Maître de Conférences (Université Lyon 1)	Co-Directeur de thèse
M. CLEMENS Francois	Professeur (TU DELFT)	Rapporteur
M. LE GAUFFRE Pascal	Maître de Conférences, HDR (INSA de Lyon)	Directeur de thèse
M. MAILHOT Alain	Professeur (INRS, Canada)	Rapporteur

Laboratoire de Génie Civil et d'Ingénierie Environnementale (LGCIE) de l'INSA de Lyon

INSA Direction de la Recherche - Ecoles Doctorales – Quinquennal 2011-2015

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	CHIMIE DE LYON http://www.edchimie-lyon.fr *Sec :Renée EL MELHEM* *Bat Blaise Pascal* *3e etage* *Insa : R. GOURDON*	M. Jean Marc LANCELIN Université de Lyon – Collège Doctoral Bât ESCPE 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04.72.43 13 95 directeur@edchimie-lyon.fr
E.E.A.	ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE http://edeea.ec-lyon.fr *Secrétariat : M.C. HAVGOUDOUKIAN* eea@ec-lyon.fr	M. Gérard SCORLETTI Ecole Centrale de Lyon 36 avenue Guy de Collongue 69134 ECULLY Tél : 04.72.18 60.97 Fax : 04 78 43 37 17 Gerard.scorletti@ec-lyon.fr
E2M2	EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION http://e2m2.universite-lyon.fr *Insa : H. CHARLES*	Mme Gudrun BORNETTE CNRS UMR 5023 LEHNA Université Claude Bernard Lyon 1 Bât Forel 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cédex Tél : 06.07.53.89.13 e2m2@ univ-lyon1.fr
EDISS	INTERDISCIPLINAIRE SCIENCES-SANTE http://www.ediss-lyon.fr *Sec :* *Insa : M. LAGARDE*	Mme Emmanuelle CANET-SOULAS INSERM U1060, CarMeN lab, Univ. Lyon 1 Bâtiment IMBL 11 avenue Jean Capelle INSA de Lyon 696621 Villeurbanne Tél : 04.72.68.49.09 Fax :04 72 68 49 16 Emmanuelle.canet@univ-lyon1.fr
INFOMATHS	INFORMATIQUE ET MATHEMATIQUES http://infomaths.univ-lyon1.fr *Sec :Renée EL MELHEM* *Bat Blaise Pascal* *3e etage* infomaths@univ-lyon1.fr	Mme Sylvie CALABRETTO LIRIS – INSA de Lyon Bat Blaise Pascal 7 avenue Jean Capelle 69622 VILLEURBANNE Cedex Tél : 04.72. 43. 80. 46 Fax 04 72 43 16 87 Sylvie.calabretto@insa-lyon.fr
Matériaux	MATERIAUX DE LYON http://ed34.universite-lyon.fr *Secrétariat : M. LABOUNE* *PM : 71.70 –Fax : 87.12* *Bat. Saint Exupéry* Ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIERE INSA de Lyon MATEIS Bâtiment Saint Exupéry 7 avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72.43 83 18 Fax 04 72 43 85 28 Jean-yves.buffiere@insa-lyon.fr
MEGA	MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE http://mega.universite-lyon.fr *Secrétariat : M. LABOUNE* *PM : 71.70 –Fax : 87.12* *Bat. Saint Exupéry* mega@insa-lyon.fr	M. Philippe BOISSE INSA de Lyon Laboratoire LAMCOS Bâtiment Jacquard 25 bis avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél :04.72 .43.71.70 Fax : 04 72 43 72 37 Philippe.boisse@insa-lyon.fr
ScSo	ScSo* http://recherche.univ-lyon2.fr/scso/ *Sec : Viviane POLSINELLI* *Brigitte DUBOIS* *Insa : J.Y. TOUSSAINT*	M. OBADIA Lionel Université Lyon 2 86 rue Pasteur 69365 LYON Cedex 07 Tél : 04.78.77.23.86 Fax : 04.37.28.04.48 Lionel.Obadia@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Tragedy is when I cut my finger. Comedy is when you fall into an open sewer and die!

Mel Brooks, *2000 year old man*

Dedication

This work is dedicated to my beloved family, specially my parents, Mehri and Mohammad for their endless support and patience, whose words of encouragement and push for tenacity have always ringed in my ears. To my sisters, Mahdieh and Elahe, two inspiration sources who have never left my side and who are very special to me. Had it not been for your help, I would never have succeeded!

Acknowledgment

First, I would like to address my most sincere gratitude and appreciations to my supervisors, Dr. Pascal Le Gauffre and Dr. Frédéric Cherqui, for their countless hours of reflection, reading, encouraging, helps and most of all patience throughout the entire thesis. It was an honor to work with you.

Second, I would also like to thank all staff members of LGCIE amongst them Pr. Sylvie Barraud, Pr. Jean-Luc Bernard-Krajewski and Dr. Gislain Lipeme Kouyi for their valuable suggestions and time.

Third, I would like to express my special thanks to *G2C Environnement*, the company in which I have carried out this work during 3 years, especially M. Christian Laplaud, M. Jean-François Closet, M. Jean-Christophe De Massiac and M. Kevin Nirsimloo for their support and patience, from whom I learnt a lot and to whom I am indebted for the opportunities they provided me.

I am also grateful for the support provided by Dr. Caty Werey, Dr. Yves Le Gat and Dr. Jean-Baptiste Aubin who guided me during the steering committees of the thesis.

I would also like to acknowledge helps provided by the Greater Lyon urban community during the thesis.

Last but not least, I tip my hat to all my friends for being with me all along the thesis from near and far; you should know that you are awesome!

This page intentionally left blank

List of publications

International peer-reviewed journals

1. **Ahmadi, M.**, Cherqui, F., Aubin, J.B., De Massiac, J.C., Le Gauffre, P. (submitted). "Sewer asset management: impact of sample size and its characteristics on the calibration outcomes of a decision-making multivariate model". *Water research Journal*.
2. **Ahmadi, M.**, Cherqui, F., De Massiac, J.C., Le Gauffre P. (2013). "From sewer inspection programs to rehabilitation needs: research and results related to data quality and availability with the support of numerical experiment*". European Journal of Civil Engineering* (In press).
3. **Ahmadi M.**, Cherqui F., De Massiac J.C., Le Gauffre P., (2013). Benefits of using basic, imprecise or uncertain data for elaborating sewer inspection programs. *Structure and Infrastructure Engineering*. (Accepted)
4. **Ahmadi M.**, Cherqui F., De Massiac J.C., Le Gauffre P., (2013). Influence of available data on efficiency of sewer inspection programs. *Urban water journal*. (DOI:10.1080/1573062X.2013.831910)
5. **Ahmadi M.**, Cherqui F., De Massiac J.C., Werey C., Lagoutte S., Le Gauffre P., (2012). "Condition grading for dysfunction indicators in sewer asset management". *Journal of Structure and Infrastructure Engineering*. (DOI: 10.1080/15732479.2012.756916).

Conference proceedings

1. **Ahmadi M**., De Massiac J.C., Cherqui F., Le Gauffre P. Nirsimloo K. (2013). "*Indigau decision support system: a new look to the sewer asset management*". The 7th IWA International Specialist conference on Efficient Use and Management of Water «water efficiency strategies for difficult times». 22-25 October, Paris, France.
2. **Ahmadi M.**, Cherqui F., De Massiac J.C., Le Gauffre P. (2013). "*Influence of availability and uncertainty of data on the efficiency of sewer inspection programs*". 5th international leading-edge conference on Strategic Asset management, LESAM, Sydney, Australia.
3. Cherqui F., **Ahmadi M.**, De Massiac J.C., Le Gauffre P. (2013). "*Impact de la disponibilité et de la qualité des données sur la prédiction d'un état de santé : expérimentation numérique*". 31st meeting of French civil engineering community. ENS Cachan, Paris, France. (http://augc2013.ens-cachan.fr/Data/Articles/Contribution1275.pdf)
4. Debères P., **Ahmadi M.**, De Massiac J.C., Le Gauffre P., Cherqui F., Werey C., Nirsimloo K. (2011). "*Deploying a sewer asset management strategy using the indigau decision support system*". 4th international leading-edge conference on Strategic Asset management, LESAM, Mulheim, Germany.
5. **Ahmadi M.**, Debères P., De Massiac J.C., Le Gauffre P., Cherqui F., Werey C. (2011). "*Using INDIGAU asset management tool for prioritizing rehabilitation programs of sewage network of Caen-la-mer*". The International Conference on Water and Wastewater, with focus on Privatization and Benchmarking, Tehran, Iran.

Abstract

Asset management is increasingly a global concern for the water and wastewater industry. It ensures that the best decisions are made for elements of the asset in order to reduce risks, optimize performance and minimize costs. Therefore, a proactive asset management which includes development of prioritization schemes for selection of inspection and rehabilitation needs is essential for a cost-effective asset management. In this manuscript, we adapt a specific definition of a proactive sewer asset management from a utility point of view going from a simple run-to-failure maintenance to a long-term proactive asset management politics. In this regard, we have identified the following bottlenecks which are addressed in this manuscript.

First the inspection surveys are currently performed based on either the judgment of the utility manager within the areas susceptible of problems (such as odors, infiltration etc.) or on reasons with a random-nature such as road works, etc. There is thus a need of elaborating inspection programs based on deterioration models in order to be more cost-effective. Despite the development of various deterioration models, the influence of availability and quality of data (imprecision, uncertainty, incompleteness) on models' predictive power has not been studied in depth. We address this issue in the first part of this thesis. We propose two methods to establish the list of the most informative factors from a representative sample of an asset stock. Among other results, this study would suggest that having imprecision in the database is preferable to having a database with no information on one specific factor. We also showed that the notion of "district" could be used instead of precise information on segment age under some hypotheses.

Second, once segments are inspected, they should be evaluated by a condition grading protocol which should remain identic for all segments. Though various condition grading protocols have been developed, they all fail to undertake the sensitivity of managers and stakeholders to the over or under-estimation of assets' condition grade as many sources of uncertainty could be found within the condition assessment process. Therefore, the need of a more specific condition grading protocol is felt which could take into account the sensitivity of utility managers and stakeholders to this specific issue. Hence, we propose a procedure in order to deal with this uncertainty. We also carry out some sensitivity analyses of parameters employed in this procedure. The results of these sensitivity tests are then applied to a part of the Greater Lyon asset stock. The results show that the assessment of segments into a condition grade depends heavily on the hypotheses that a manager could make about the under or over-estimation of its assets' condition.

Third, at the moment small number of utilities has completely inspected and evaluated their asset stocks. Therefore, the use of a representative sample from an asset stock in order to calibrate decision-support models as deterioration models seems mandatory. Nevertheless, in this regard we should tackle with following problematic issues: 1) How to draw a representative sample of an asset stock which reflects the characteristics of this asset stock? and 2) What is the impact of used sample on the calibration outcomes of these multivariate models? Hence, by drawing several samples with different sizes according to different sampling methods and applying Monte Carlo method, we have proposed a procedure in order to study the influence of available sample on the outcomes of a multivariate model. By proposing some statistical analyses, we showed that the calibration process depends extremely on available sample which could results, if this latter is not drawn properly, into erroneous conclusions.

Keywords: Asset management, sewer network, decision support, sewer inspection program, condition grade, logistic regression.

Résumé

La gestion patrimoniale est une problématique d'importance croissante pour les gestionnaires des réseaux d'eau potable et d'assainissement. Elle vise à choisir les meilleures décisions d'actions à mener sur les éléments du patrimoine, pour limiter les risques, optimiser les performances et réduire les couts. Cela implique une démarche proactive incluant le développement d'outils de hiérarchisation pour sélectionner les conduites à inspecter et/ou réhabiliter. Dans ce manuscrit, nous avons adapté la définition spécifique de la gestion proactive des réseaux d'assainissement aux pratiques des gestionnaires pour passer de la maintenance post-défaillance à une politique de gestion patrimoniale proactive sur le long-terme. Pour cela, nous avons identifié les verrous suivants qui sont abordés dans le manuscrit.

Premièrement, les inspections des conduites sont actuellement programmées soit selon le jugement du gestionnaire ciblant les secteurs problématiques (odeurs, infiltrations, etc.), soit indépendamment des choix du gestionnaire (lors des travaux de voirie programmés, etc.). Il est donc nécessaire de pouvoir élaborer ces programmes d'inspections à partir de modèles de détérioration pour optimiser les budgets alloués. Malgré le développement de nombreux modèles de détérioration, l'influence de la qualité et de la disponibilité des données (incomplétude, imprécision, incertitude) sur la puissance prédictive des modèles n'a pas été étudiée en détail. Nous avons abordé cette question dans la première partie de la thèse, en proposant deux méthodes pour déterminer la liste des facteurs les plus informatifs à partir d'un échantillon représentatif. Entre autres, nos résultats suggèrent que l'imprécision sur une donnée est préférable à ne pas disposer de cette donnée. Nous avons également montré que la notion de « quartier » pourrait être utilisée, sous certaines conditions, pour compenser la non-connaissance de l'âge des conduites.

Deuxièmement, une fois les conduites inspectées, leur état de santé doit être évalué à l'aide d'un protocole identique pour l'ensemble du patrimoine. Bien que différents protocoles existent, aucun ne permet de prendre en compte les considérations du gestionnaire concernant la sur- ou sous-estimation de l'état de santé (liée aux nombreuses incertitudes induites par l'ensemble du processus). Il est donc nécessaire que le protocole utilisé puisse minimiser les erreurs de sur- ou sous-estimation en fonction des choix des gestionnaires. Nous avons proposé une procédure prenant en compte ces choix, ainsi que l'état global du patrimoine considéré. Les études de sensibilité réalisées à partir des données d'inspection d'une partie du patrimoine du Grand Lyon montrent que l'état de santé évalué dépend fortement des choix faits par le gestionnaire concernant la sur- ou sous-estimation.

Troisièmement, peu de patrimoines ont actuellement été complètement inspectés et évalués. Ainsi, l'utilisation d'un échantillon représentatif d'un patrimoine semble obligatoire pour pouvoir calibrer des modèles justifiant les décisions comme par exemple un modèle de détérioration. Néanmoins, cela pose les problèmes suivants : 1) comment générer un échantillon représentatif reflétant au mieux les caractéristiques du patrimoine complet ? 2) Quel est l'impact de cet échantillon sur les résultats de calage des modèles multi-variables ? Ainsi, après avoir généré différents échantillons de différentes tailles et selon différentes méthodes d'échantillonnage, à l'aide de la méthode de Monte Carlo, nous avons pu proposer une procédure pour étudier l'influence des caractéristiques de l'échantillon sur les résultats du modèle de détérioration. A partir d'analyses statistiques, nous avons montré que le processus de calage (des modèles) dépend fortement de l'échantillon disponible, et cela peut donc conduire à des résultats erronés.

Mots clefs : gestion patrimoniale, assainissement, aide à la décision, programme d'inspection des conduites, état de santé, régression logistique.

This page intentionally left blank

Contents

Chapitre I: Synthèse du mémoire en français

1.1. Contexte et questions formulées

Les réseaux d'assainissement font parties des infrastructures à forte valeur économique dans les villes d'aujourd'hui. Pour préserver leur bon fonctionnement au cours du temps, ils doivent donc être gérés de manière efficace. Traditionnellement, les collectivités ont abordé la conception, la construction, l'entretien et l'exploitation de ces infrastructures selon une approche réactive, basée sur les éventuels problèmes constatés sur le patrimoine (Wirahadikusumah *et al.*, 2001).

La mise en place d'une réelle gestion patrimoniale des réseaux nécessite le développement d'une approche proactive comprenant la programmation de l'inspection du patrimoine, l'évaluation de l'état et des performances du patrimoine, l'estimation des besoins en réhabilitation, et la priorisation des actions compte tenu des contraintes budgétaires.

Nous fournissons une définition spécifique de cette approche proactive en nous adaptant au point de vue des collectivités (figure 1-1 adaptée de Ahmadi *et al.*, 2013). La figure 1-1 montre les étapes clefs de cette définition qui, en tout, contient 4 niveaux progressifs de gestion patrimoniale à mettre en œuvre par les collectivités. Ces niveaux varient d'une simple approche de fin de vie des tronçons à un programme proactif long-terme.

Ces niveaux sont les suivants :

(1) Définition des besoins en matière de réhabilitation en évaluant l'état de santé des tronçons inspectés ;

(2) Priorisation des besoins en matière de réhabilitation en combinant l'évaluation de l'état de santé des tronçons avec la connaissance de dysfonctionnements observés

et/ou avec l'évaluation de la vulnérabilité de l'environnement des tronçons (ces deux niveaux sont limités aux tronçons inspectés) ;

(3) Priorisation des besoins en matière d'inspection en considérant une ou la combinaison des informations suivantes :

(a) Vulnérabilité de l'environnement des tronçons ;

(b) Etat supposé des tronçons non investigués ;

(4) La planification des inspections/ réhabilitations et l'élaboration de plans budgétaires à long-terme à partir d'un échantillon représentatif du patrimoine considéré ;

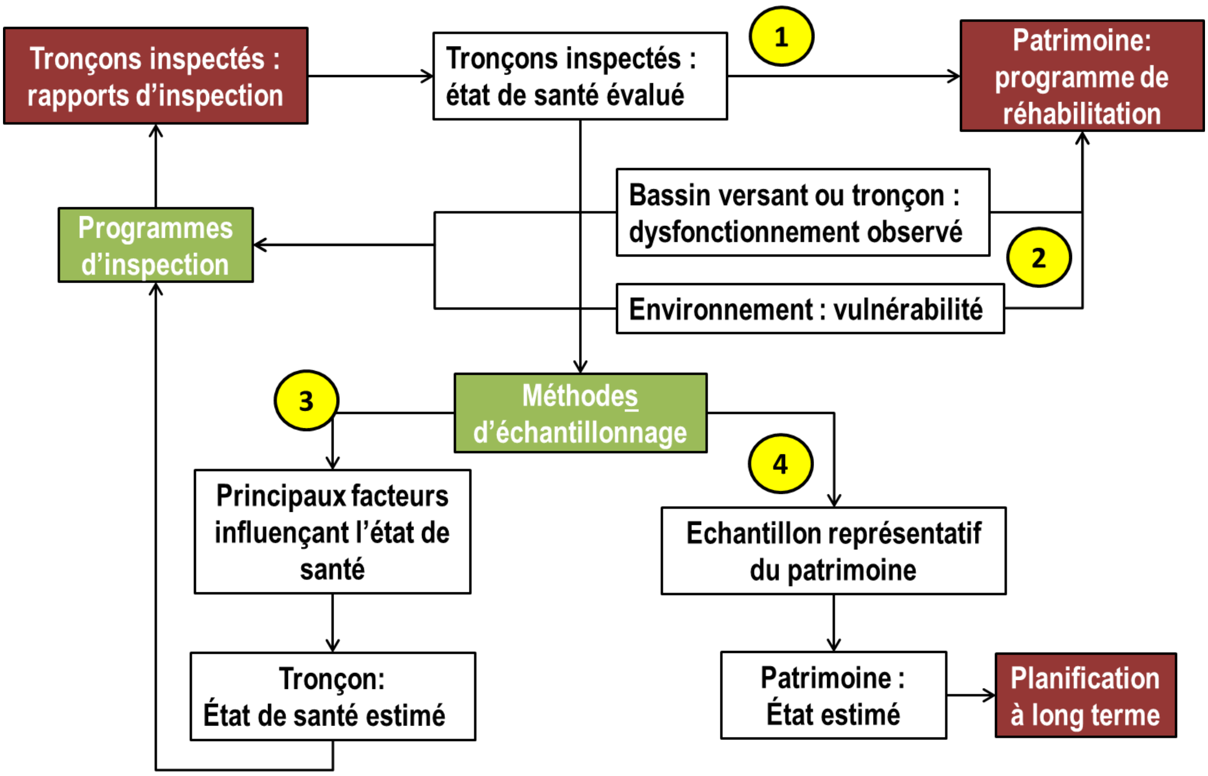


Figure 1-1. Les niveaux de la gestion patrimoniale proactive des réseaux d'assainissement

En France, les premiers et deuxièmes niveaux ont été développés au cours des projets français RERAU (Réhabilitation des réseaux d'assainissement urbains) et INDIGAU (Indicateurs de performance pour la gestion patrimoniale des réseaux d'assainissement urbains) et qui sont basés sur les ITVs (inspections télévisées) conformes à la norme Européenne EN 15308-2. Des indicateurs de dysfonctionnement construits à partir des codes inventaires prévus dans la norme Européenne sont fournis dans ces projets. Trois outils destinés à l'usage direct des gestionnaires ont également été développés. Ces travaux ont été élaborés entre 2001 et 2010 (Le Gauffre *et al.*, 2004; Le Gauffre *et al.*, 2007; Le Gauffre *et al.*, 2010).

1.1.1. Questions relatives à la programmation des inspections

Ces deux premiers niveaux de sophistication de la gestion patrimoniale (figure 1-1) sont limités à des tronçons déjà inspectés. Autrement dit, tous les tronçons non-inspectés sont exclus de la procédure de prise de décision par rapport aux besoins en matière de réhabilitation. D'autre part, les campagnes d'inspections sont normalement basées soit sur le jugement des gestionnaires du réseau (par exemple identifier d'éventuels problèmes d'odeur, d'infiltration etc.) soit sur des motifs qui ont une nature aléatoire comme des travaux de voirie par exemple. Par conséquent, pour améliorer la démarche de gestion patrimoniale, il est nécessaire d'élaborer des programmes d'inspection en lien avec l'état de santé supposé des tronçons non inspectés. Cette nécessité est forte, à cause du grand nombre de tronçons problématiques résultant des activités d'entretien non-systématiques, ainsi qu'à cause des restrictions budgétaires. Pourtant, peu d'approches opérationnelles sont actuellement disponibles en tant qu'outils d'aide à la décision pour planifier des compagnes d'inspection (Baur et Herz 2002).

Afin d'évaluer le risque de défaillance d'un tronçon, sa probabilité de défaillance doit être combinée avec les conséquences de sa défaillance (Le Gauffre *et al.*, 2007). La probabilité de défaillance peut être obtenue à partir des modèles de détérioration. Les outils destinés à prioriser des besoins en matière d'inspection devraient être fondés sur ces modèles de détérioration qui permettent de prévoir l'état actuel ou futur des tronçons à partir de la connaissance des facteurs explicatifs.

Bien que différents types de modèles de détérioration peuvent être trouvés dans la littérature scientifique, l'attention se concentre pour l'instant toujours sur le type de modèle de détérioration utilisé. Ainsi l'influence de la qualité et de la disponibilité des données (en termes d'incertitude, d'imprécision et d'incomplétude) sur la puissance prédictive de ces modèles n'a pas été étudiée. Les questions principales que nous abordons concernant la qualité des données sont donc les suivantes :

(1) Quels facteurs explicatifs sont les plus informatifs en termes de programmes d'inspection? Quelles données faut-il privilégier pour prédire l'état de santé des tronçons ? Peut-on établir la liste des facteurs les plus informatifs ?

(2) Quel est le gain apporté par des données imprécises en comparaison à une absence de ces données (imprécision vs incomplétude) ?

(3) Les données les plus probablement disponibles dans une collectivité (par exemple les données recueillies pour l'utilisation des modèles hydrauliques, telles que le diamètre, la profondeur, la pente, le type de réseau et la longueur) peuvent-elles être suffisantes pour définir un programme d'inspection efficace ?

(4) Pouvons-nous utiliser une variable auxiliaire afin de compenser les effets des données manquantes ?

(5) Est-il préférable d'accepter un certain degré d'incertitude dans les données au lieu de ne pas avoir ces données ?

1.1.2. Questions relatives à l'évaluation d'un état de santé à partir de l'inspection

En outre, une fois que les tronçons sont inspectés, ils doivent être évalués par un protocole d'évaluation de l'état de santé qui restera identique pour tous les tronçons. Bien que différents protocoles aient été développés tels que WRc, NAAPI, PACP etc., tous ne parviennent pas à considérer le défi suivant qui a été l'objet d'une partie de cette thèse :

- Le processus d'évaluation de l'état de santé des tronçons contient des sources d'incertitude. Ce processus peut entraîner une surestimation ou une sous-estimation de l'état de santé réel du tronçon. Toutefois, les protocoles existants ne permettent pas aux gestionnaires de prendre en compte leurs propres considérations sur la surestimation ou la sous-estimation de l'état de santé de leurs patrimoines. Par conséquent, la nécessité d'un protocole qui peut éventuellement prendre en compte la sensibilité des gestionnaires vis-à-vis de cette question spécifique se fait sentir.

1.1.3. Questions relatives à l'étude du patrimoine à partir d'un échantillon

Aujourd'hui, peu de collectivités ont complètement inspecté et évalué leurs patrimoines. L'usage d'un échantillon représentatif d'un patrimoine semble donc obligatoire afin de calibrer des modèles d'aide à la décision comme des modèles de détérioration ainsi d'étudier des scénarios de futur pour en déduire des conclusions appropriées qui portent sur l'ensemble du patrimoine et pas seulement la partie inspectée. A cet égard, nous formulons les questions suivantes :

- Comment constituer un échantillon représentatif d'un patrimoine qui montre de manière correcte des caractéristiques souhaitées de ce patrimoine ?
- Comment fournir une estimation fiable d'une caractéristique du patrimoine étudié à partir de cet échantillon dit représentatif?
- D'autre part, le calage des modèles multi-variables est délicat. Quels sont des impacts de l'échantillon utilisé sur des résultats du calage de ces modèles ?

Pour résumé, les objectifs de cette thèse sont donc les suivants:

- Proposer une approche systématique pour les programmes d'inspection et par la suite, étudier l'influence de la qualité et la disponibilité des données sur ces derniers ;
- Vérifier la sensibilité du protocole d'évaluation de l'état de santé proposé, qui avait été développé durant les projets français RERAU et INDIGAU, en termes du paramétrage requis ;

- Prélever un échantillon représentatif d'un patrimoine qui montre de manière correcte des caractéristiques de ce patrimoine ;
- Etudier l'impact de l'échantillon utilisé sur des résultats du calage des modèles multi-variables.

Ces objectifs sont traités dans les trois parties principales de cette thèse :

(1) Elaborer des programmes d'inspection et étudier l'impact de la qualité et la disponibilité des données ;

(2) Evaluer un tronçon à partir d'une inspection ;

(3) Evaluer un patrimoine à partir d'un échantillon.

1.2. Elaboration des programmes d'inspection et étude de l'impact de la qualité et de la disponibilité des données

L'un des objectifs de cette partie de la thèse est de proposer une approche systématique afin d'élaborer des programmes d'inspection en utilisant un modèle de détérioration (*Segment targeting model* sur la figure 1-2). L'utilisation d'un modèle de détérioration doit permettre de réaliser des programmes d'inspection plus efficients, c'est-à-dire d'augmenter leur rendement : nombre de tronçons en mauvais état identifiés grâce à ces inspections. L'autre objectif est d'étudier l'influence des données disponibles dans une collectivité sur des programmes d'inspection. A cet effet, une base de données complète doit être disponible. Dans une démarche d'expérimentation numérique, l'information que contient cette base de données sera dégradée en introduisant de l'incomplétude, de l'imprécision et de l'incertitude de manière maîtrisée pour représenter la réalité des bases de données des collectivités (fig. 2).

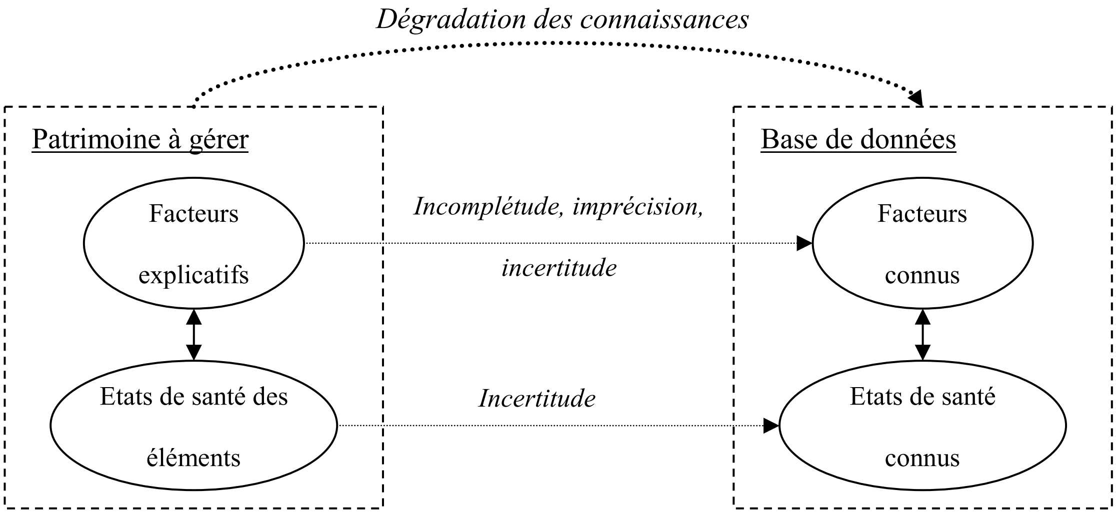


Figure 1-2. Différence entre le patrimoine complet et la base de données d'une collectivité.

Notre démarche d'expérimentation numérique applique et évalue une approche systématique de programmation des inspections basée sur un modèle de détérioration (figure 1-3) et sur les hypothèses exposées ci-dessous. Nous supposons qu'avant d'élaborer un nouveau programme d'inspection, une collectivité avait déjà inspecté 10% de la longueur totale de son patrimoine. En outre, nous supposons qu'elle inspecte, chaque année, 5% de la longueur totale de son patrimoine. Ce dernier peut se diviser en deux parties : 1- les tronçons ciblés par les programmes d'inspection (2% de la longueur totale du patrimoine) et 2- les tronçons inspectés de manière aléatoire (3% de la longueur totale du patrimoine inspectés pour différentes raisons comme des travaux de voirie, etc.).

Le choix des tronçons ciblés est réalisé en utilisant un modèle de détérioration de type régression logistique qui permet d'évaluer la probabilité d'être en mauvais état pour les tronçons non inspectés (un état critique que nous allons définir dans la partie suivante de la thèse). Nous faisons l'hypothèse que ce programme s'effectue pendant 4 années.

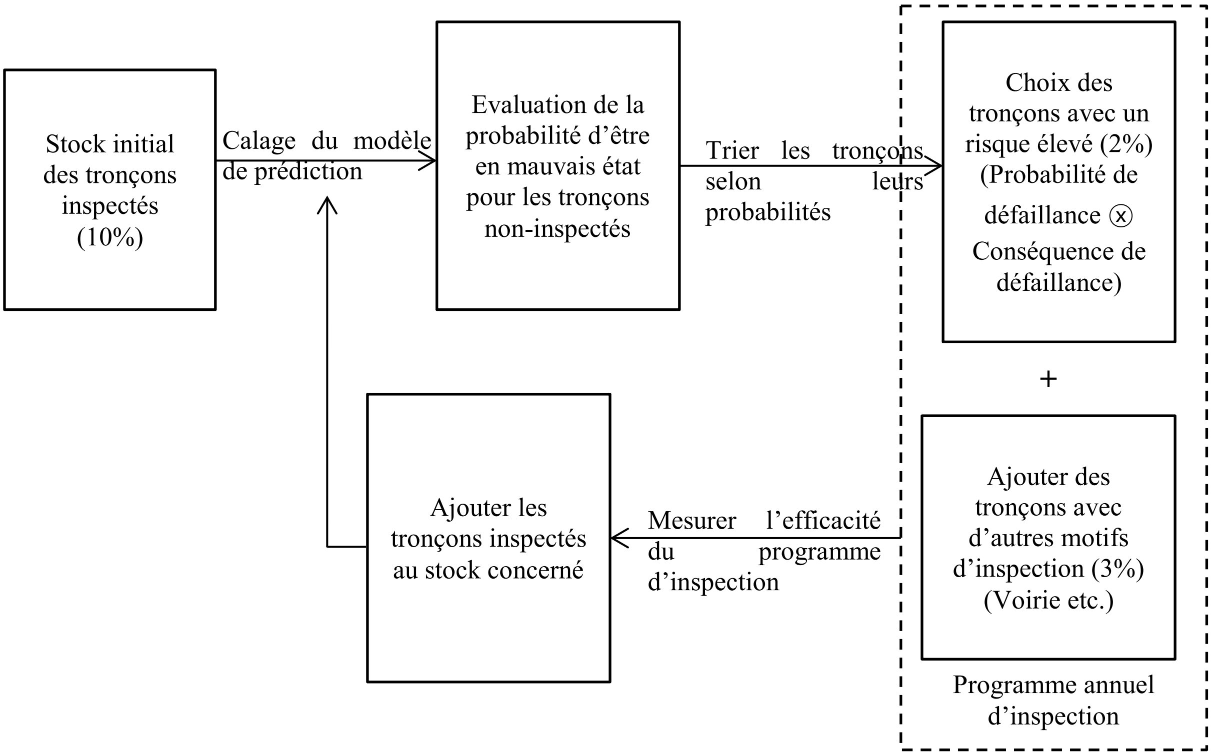


Figure 1-3: Expérimentation numérique d'une approche systématique pour les programmes d'inspection

En appliquant cette approche systématique pour effectuer des expérimentations numériques, nous proposons quelques indicateurs pour mesurer l'efficacité de chaque scénario définis dans le tableau 1-1. Ce tableau récapitule aussi les questions que nous nous sommes posées dans les paragraphes ci-dessus.

Tableau 1-1: Les questions concernant la qualité et la disponibilité des données

Questions	Étudier l'influence de
1- Quels facteurs explicatifs sont les plus informatifs ? Peut-on établir la liste des facteurs les plus informatifs ?	L'incomplétude dans les données
2- L'incomplétude peut-elle être compensée ? Peut-on pallier la non-connaissance d'un facteur prépondérant par d'autres facteurs	L'incomplétude et l'imprécision dans les données
3- Comment les données les plus probablement disponibles dans une collectivité peuvent être utilisées pour définir un programme d'inspection efficace? Pouvons-nous utiliser une variable auxiliaire afin de compenser les effets des données manquantes?	L'incomplétude et l'imprécision dans les données
4- Est-il mieux d'accepter un certain degré d'incertitude dans les données au lieu de ne pas les avoir?	L'incomplétude et l'incertitude dans les données

En considérant une base de données semi-virtuelle (figure 1-5) qui contient 8 facteurs explicatifs de l'état des tronçons ainsi que leurs états de santé, nous avons étudié l'influence de l'incomplétude, de l'imprécision et de l'incertitude dans les données disponibles sur l'efficience des programmes d'inspection. Nous avons montré que la qualité des données en termes d'imprécision, d'incomplétude et d'incertitude a une influence majeure sur l'efficacité des programmes d'inspections. Nous montrons que l'imprécision et/ou l'incertitude dans les données sont préférables à l'absence de ces données (l'incomplétude). Toutefois, si quelques données primordiales ne sont pas disponibles, nous pouvons envisager d'utiliser des variables auxiliaires afin de compenser des effets de leurs absences dans la base de données.

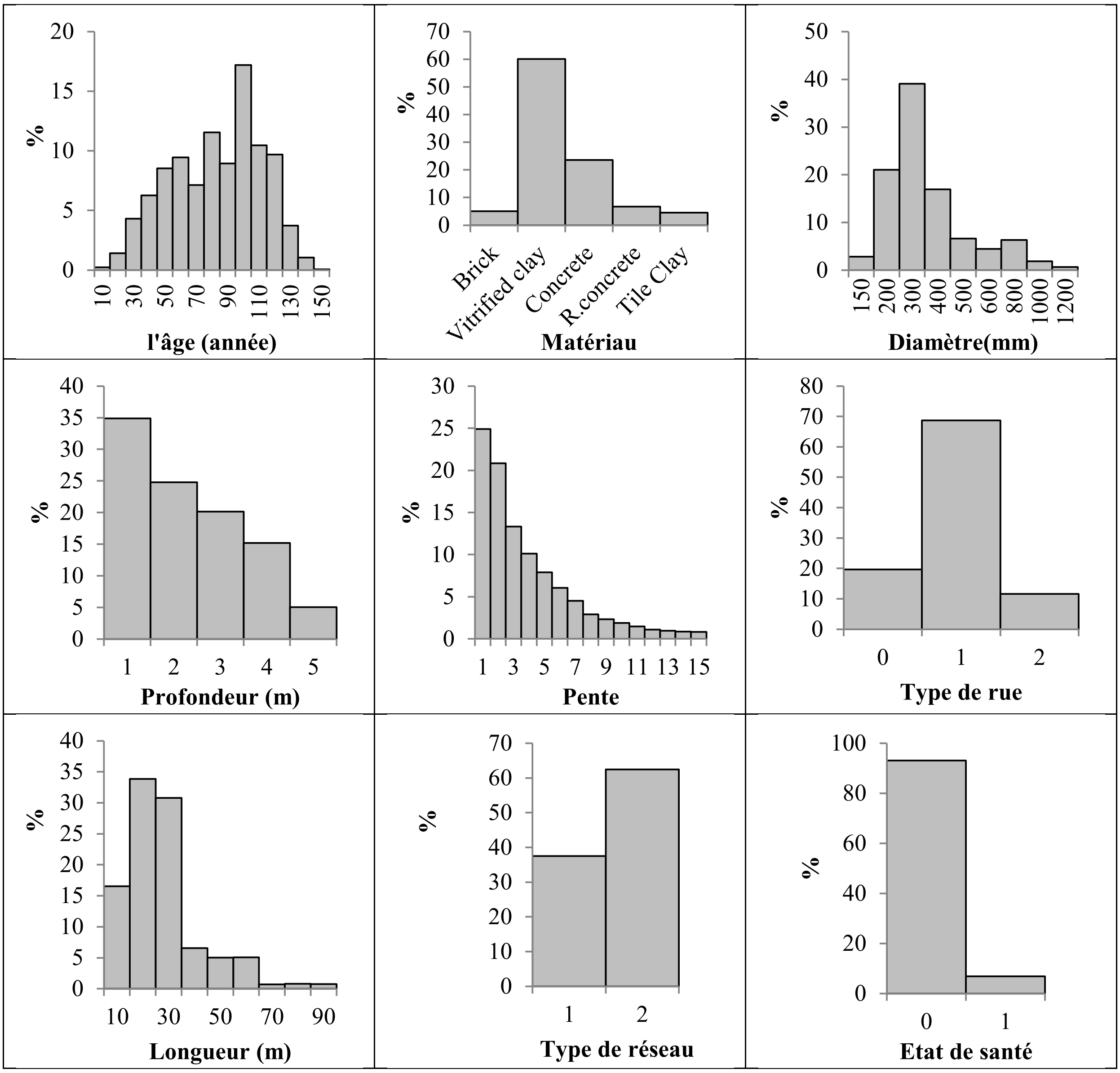


Figure 1-5: l'histogramme des facteurs explicatifs et l'état de santé des tronçons

Nous avons aussi pu élaborer la liste des facteurs les plus informatifs parmi les facteurs disponibles en proposant 2 approches différentes :

- Une méthode d'analyse statistique basée sur la notion de *Déviance* qui est un test de maximum de vraisemblance, plus abstraite pour un gestionnaire, mais applicable sur tout échantillon du patrimoine. Cette analyse statistique utilise la méthode d'élimination progressive (*backward selection procedure*) des variables indépendantes ;
- La simulation de plusieurs scénarios de collecte de données par le gestionnaire et l'évaluation du gain : capacité à atteindre les tronçons défaillants lors des inspections ;

Les résultats de ces deux approches sont identiques. Cela permet d'appliquer la première approche sur un échantillon représentatif du patrimoine étudié afin d'établir la liste des facteurs les plus informatifs et qui nous permettra, par la suite, de construire des plans d'acquisition de données.

1.3. Evaluation d'un tronçon à partir d'une inspection

Lorsqu'un tronçon est inspecté, il doit être évalué. Alors que les inspections télévisées (ITV) sont largement utilisées pour évaluer l'état des tronçons, de nombreuses sources d'incertitude se trouvent dans le processus d'élaboration des programmes de réhabilitation à partir des résultats des compagnes d'inspection.

Nous avons identifié 4 sources principales d'incertitude dans le processus d'évaluation:

- La procédure de saisie des codes inventaires (défauts) existants sur un tronçon selon une norme d'inspection telle que PACP, EN 13508-2 ou WRc par un vidéaste ;
- La conversion de ces codes inventaires en une note synthétique pour chaque tronçon. Cette quantification dépend de l'extension, de la sévérité et de la forme des défauts.

- L'affectation d'un tronçon à un état de santé en considérant sa note synthétique. Cet aspect nécessite la définition et l'utilisation de seuils d'affectation qui peut donc être l'une des sources d'incertitude.
- Le croisement d'un indicateur issu des inspections télévisées tel que *Infiltration* avec d'autres indicateurs issues des données diverses comme la vulnérabilité de l'environnement afin de constituer un critère de réhabilitation.

Comme nous l'avons dit dans la section 1.1, les protocoles existants ne permettent pas aux gestionnaires de prendre en compte leurs propres considérations sur la surestimation ou la sous-estimation de l'état de santé de leurs tronçons. Par conséquent, la nécessité d'un protocole qui peut éventuellement prendre en compte la sensibilité des gestionnaires vis-à-vis de cette question spécifique se fait sentir.

Un protocole a été développé dans les projets français RERAU et INDIGAU qui est basé sur un échantillon de tronçons de différentes collectivités françaises. Ces tronçons ont été évalués par plusieurs gestionnaires français pour tous les indicateurs de dysfonctionnement qui ont été définis durant ces projets (tableau 1-2) et évaluables à l'échelle d'un tronçon.

Tableau 1-2. Les indicateurs de dysfonctionnement selon la guide méthodologique RERAU

Indicateur	Définition
ABR	Abrasion
BOU	Bouchage
EFF	Effondrement : altération de l'intégralité structurale
COR	Corrosion
DEB	Débordement
EXF	Exfiltration
ATC	Attaque chimique
HYD	Diminution de la capacité hydraulique
INF	Infiltration
RAC	Dégradation par intrusion des racines
ENS	Ensablement
DSC	Déstabilisation du complexe sol-conduite

En outre, ces projets proposent un tableau de quantification pour chaque indicateur de dysfonctionnement (par exemple pour *Infiltration :* tableau 1-3). Chaque défaut est traduit, à l'aide de ces tableaux, en une note élémentaire qui tient compte de la sévérité et de l'extension du défaut. L'extension est soit la longueur du défaut continu soit une valeur prédéterminé *b* qui est attribuée pour des défauts ponctuels. Le degré de sévérité de chaque défaut est décrit à l'aide d'un seul paramètre *a* et de 4 niveaux de gravité : mineur (1), moyen (*a*), grave (a^2) et très grave (a^3). Ainsi, chaque défaut (non majeur) observé est traduit en une note N_i :

$$N_i = a^n \times b(ou\ L_i), \text{avec}\ n = 0, 1, 2\ \ ou\ 3\ \text{et}\ a = 2, 3\ \ ou\ 4$$

Tableau 1-3. Quantification de INF (Le Gauffre *et al.*, 2004).

Observation Oi	Code Ci	1	a	a^2	a^3	← Sévérité Extension ↓
Déformation	BAA		BAA			b
Fissure	BAB	BAB B		BAB C		Li
Rupture/ Effondrement	BAC			BAC A	BAC B/C	b
Mortier manquant	BAE		BAE			b
Raccordement défectueux	BAH			BAH B/C/D		b
…	…					

Ensuite, la note moyenne attribuée à chaque tronçon est la somme de ces notes élémentaires divisée par la longueur du tronçon (comme pour les notes WRc et NRC).

Tous les tronçons évalués par des experts de différentes collectivités françaises sont aussi calculés par cette procédure de quantification. Ensuite, un critère de calage (coût des erreurs) permet de fixer les seuils nécessaires pour affecter des tronçons à des états de santé de G1 (situation acceptable) à G4 (situation intolérable). La figure 1-6 montre la distribution des évaluations d'expert versus les notes calculées des tronçons pour l'indicateur *Infiltration*.

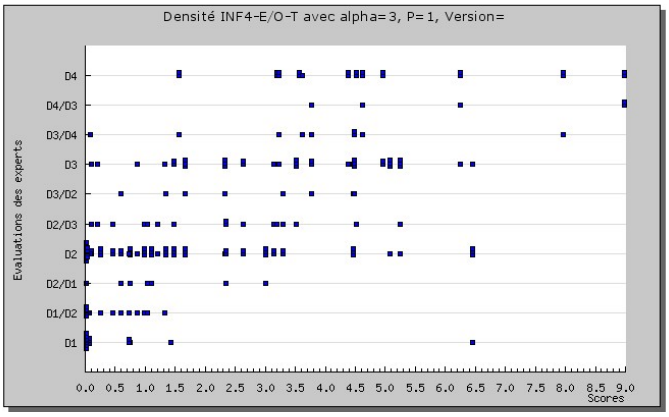


Figure 1-6. Distribution des avis d'expert versus la note calculée

Le critère de calage proposé est basé sur deux ensembles de paramètres:

- L'état global du patrimoine étudié, à travers quatre probabilités $P(E_i)$;
- Les coûts ou poids (w_{ij}) des différents types d'erreurs, qui représentent la sensibilité des gestionnaires à la surestimation ou la sous-estimation de l'état de dégradation des tronçons.

Nous avons effectué une analyse de sensibilité de ces paramètres. Le tableau 1-4 et la figure 1-7 décrivent ces jeux de sensibilité. Nous avons montré que tous les deux paramètres ont une influence sur les seuils d'affectation (figure 1-8). Les résultats de ces jeux de sensibilité ont été appliqués sur 4471 tronçons du patrimoine du Grand Lyon (tableau 1-5).

Tableau 1-4. Trois hypothèses sur l'état global du patrimoine étudié : P1, P2, P3 sont les probabilités *$P(E_i)$* qu'un tronçon est dans l'état G_i

	G1	G2	G3	G4
P1	10%	40%	40%	10%
P2	30%	30%	20%	20%
P3	20%	30%	30%	20%

M1

0	1	3	5
1	0	1	3
3	1	0	1
5	3	1	0

M2

0	1	1	1
1	0	1	1
5	2	0	1
10	5	3	0

M3

0	1	1	1
3	0	1	1
15	6	0	1
30	15	9	0

Figure 1-7. Trois hypothèses sur les coûts des erreurs d'affectation (w_{ij}). En vertical, de G1 à G4 évalué par Expert, en horizontale, de G1 à G4 évalué par le calcul

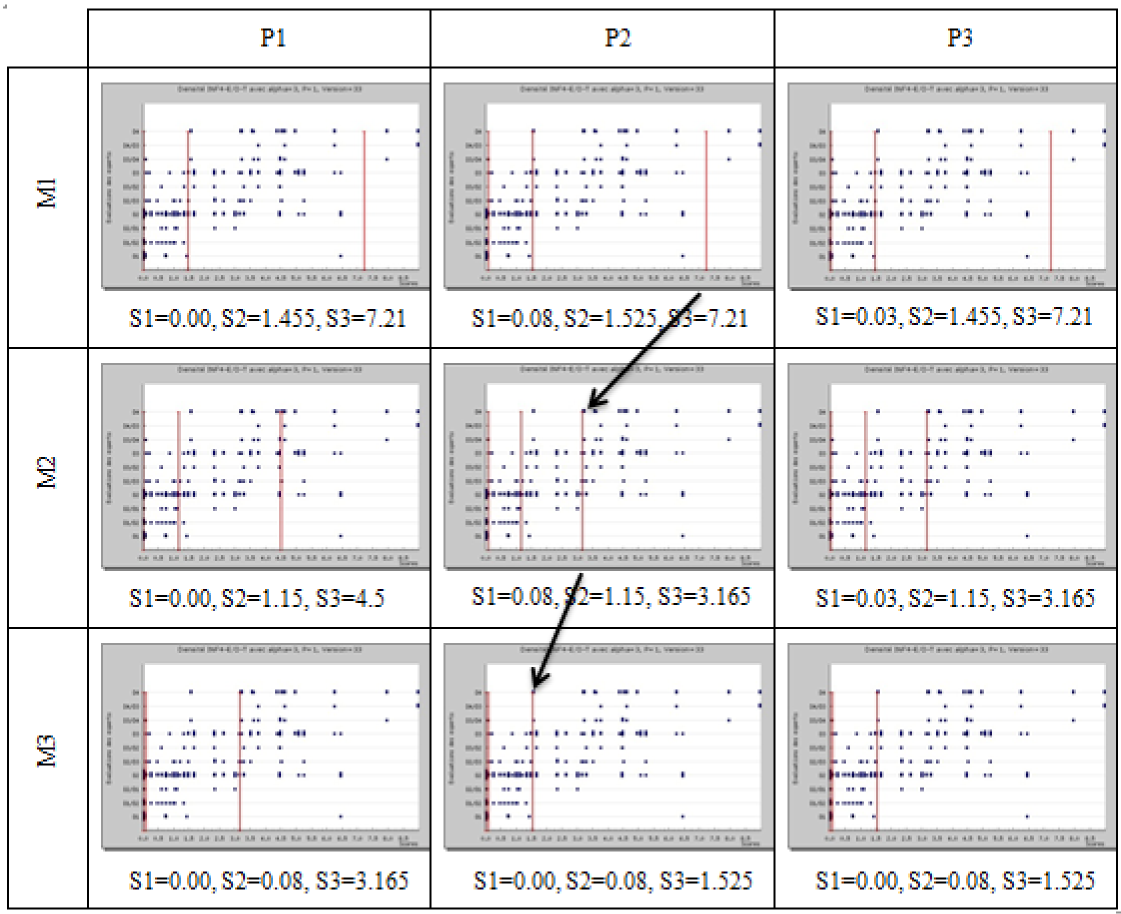


Figure 1-8. Résultats de l'analyse de sensibilité

Tableau 1-5: évaluation d'une partie du patrimoine du Grand Lyon en pourcentage de G1 à G4

		P1	P2	P3
M1	G1	46%	52%	48%
	G2	42%	37%	40%
	G3	9%	8%	9%
	G4	3%	3%	3%
M2	G1	46%	52%	48%
	G2	39%	33%	37%
	G3	11%	10%	10%
	G4	4%	5%	5%
M3	G1	46%	46%	46%
	G2	6%	6%	6%
	G3	43%	37%	37%
	G4	5%	11%	11%

Après avoir appliqué les résultats de l'analyse de sensibilité sur 4471 tronçons du Grand Lyon, pour l'indicateur *Infiltration*, nous constatons que les six premiers jeux de paramètres (scénarios) fournissent plus ou moins le même résultat. Autrement dit, presque 85% et 15%

des tronçons sont respectivement en G1-G2 et G3-G4. Cela montre que les résultats ne varient pas de manière significative d'un jeu à l'autre alors que presque tous les seuils varient. La variation du pourcentage des tronçons évalués en G3-G4 en passant d'un jeu à l'autre est de l'ordre de 4%. Par contre, un plus grand nombre de tronçons sont évalués en G3-G4 avec les trois derniers jeux de paramètres (48%). Cela est dû au fait que pour ces scénarios nous utilisons une matrice de coûts des erreurs qui est plus sécuritaire : induisant moins de faux négatifs mais donc plus de faux positifs.

1.4. Evaluation d'un patrimoine à partir d'un échantillon

Les objectifs principaux de cette partie de la thèse sont les suivants:

- Comment générer un échantillon représentatif d'un patrimoine qui montre de manière correcte des caractéristiques de ce patrimoine ?
- Comment fournir une estimation fiable d'une caractéristique du patrimoine étudié à partir de cet échantillon dit représentatif?
- D'autre part, le calage des modèles multi-variables est délicat. Quels sont les impacts de l'échantillon utilisé sur les résultats du calage de ces modèles ?

Pour répondre à la première question, nous avons considéré 3 méthodes d'échantillonnage : aléatoire simple (*SRS : simple random sampling*) et les allocations proportionnelle et de Neyman (optimum) dans la méthode stratifiée (*PSRS : proportional allocation in stratified sampling & OSRS : Optimum allocation in stratified sampling*).

En outre, pour répondre à la deuxième et à la troisième questions, nous supposons que n est la taille de l'échantillon disponible pour 'caler' un modèle multi-variables, ou plus précisément pour réaliser l'estimation des paramètres de ce modèle. Nous choisissons la régression logistique pour être cohérent avec d'autres parties de la thèse. Nous faisons varier n de 600 à 5000 tronçons avec un pas de 400 tronçons. En appliquant la méthode de Monte Carlo, nous effectuons 1000 essais (1000 échantillons différents) pour chaque valeur d'effectif *n.* Puis, pour chaque échantillon généré, nous estimons les paramètres de la régression logistique en appliquant la méthode de maximum de vraisemblance (Agresti 2002). Les résultats trouvés, les coefficients et leurs écarts-type, sont enregistrés pour une analyse plus approfondie. Nous nous fixons un seuil de convergence égale à 10^{-5} pour la méthode du maximum de vraisemblance. Il est important de noter que nous avons réalisé cette procédure pour les trois méthodes d'échantillonnage considérées.

Les résultats des simulations numériques pour les échantillons de différentes tailles sont évalués relativement aux résultats du modèle complet qui est basé sur la population mère. L'exactitude (*accuracy* en anglais), la précision et la signification statistique de chaque coefficient sont aussi évaluées. Ces évaluations s'effectuent en définissant quelques indicateurs statistiques. Tous les indicateurs sont calculés sous la condition de la convergence de la régression logistique. Les échantillons pour lesquels la régression logistique n'obtient pas la convergence sont exclus de notre analyse.

La figure 1-9 montre le patrimoine utilisé pour les simulations numériques. Ce patrimoine contient les facteurs explicatifs suivants pour chaque tronçon : l'âge, le matériau, le type de réseau, la pente, le type de rue, la profondeur, le diamètre, la longueur et l'état de santé. En tout, il contient 9810 tronçons dont la longueur totale est égale à 213 km.

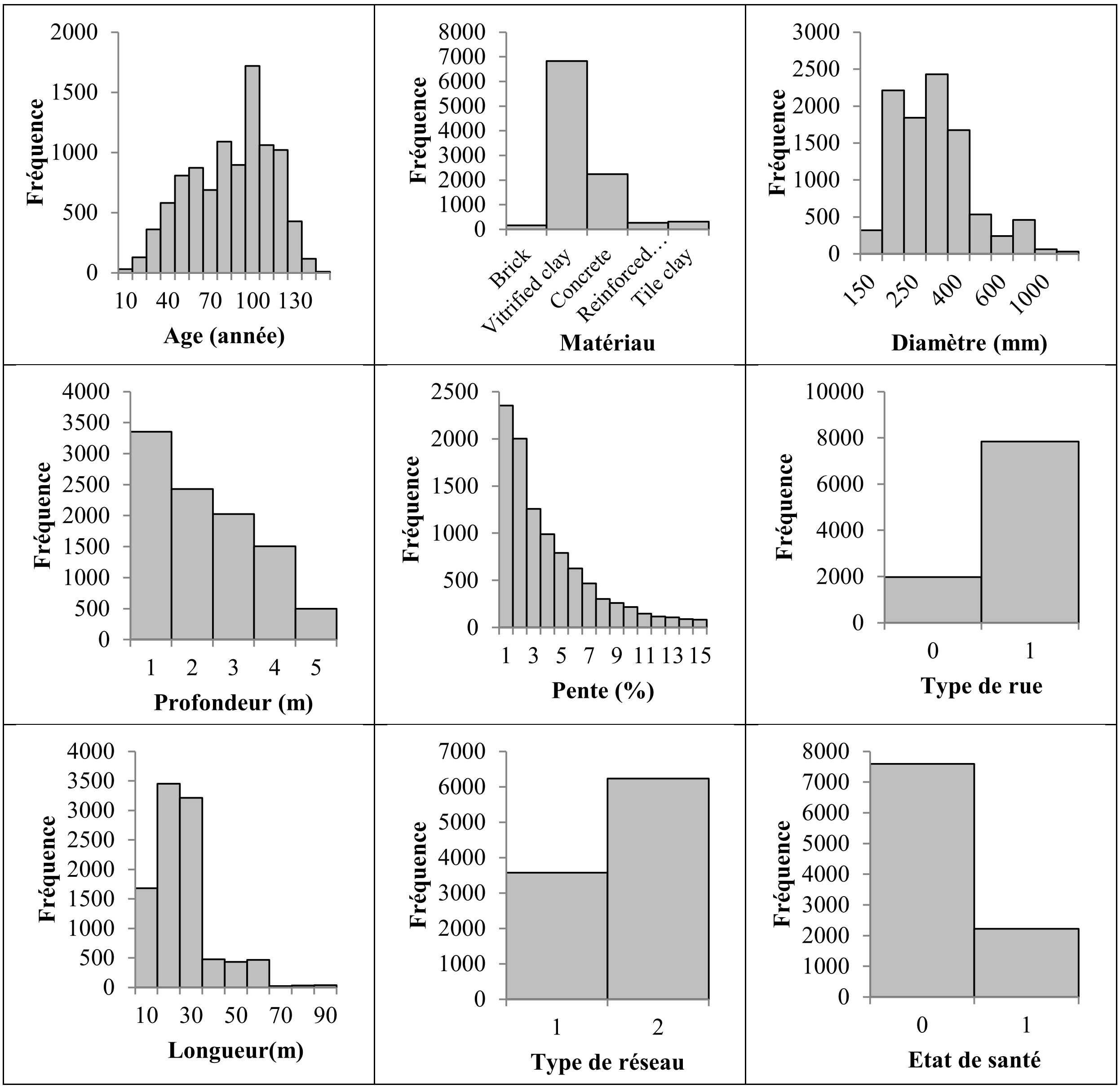


Figure 1-9: histogrammes des facteurs explicatifs présumés et de l'état de santé des tronçons

La première phase de l'étude concerne la notion d'échantillon représentatif pour l'étude de la variable « pourcentage de tronçons en mauvais état ». Les résultats des simulations numériques montrent qu'il n'y a pas de différences significatives, dans notre cas, entre les différentes méthodes d'échantillonnage en termes de taille requise de l'échantillon. Les résultats montrent qu'avec un échantillon dont l'effectif est de 1000 tronçons, toutes les méthodes d'échantillonnage choisies fournissent un échantillon représentatif du patrimoine. En outre, le tableau 1-6 montre les pourcentages de simulations dont la valeur estimée est entre la valeur vraie sur la population (la proportion des tronçons en mauvais état) ± 3%.

Tableau 1-6: % de simulations dont la valeur estimée est entre la valeur vraie sur de la population ± 0.03

n	Aléatoire simple	Allocation proportionnelle Méthode stratifié	Allocation de Neyman (Optimum) ; méthode stratifié
600	92.4	91.7	93.3
1000	98.6	97.9	98.9
1400	99.8	99.5	99.6
1800	99.9	99.9	99.9
2200	100	100	100
2600	100	100	100
3000	100	100	100
3400	100	100	100
3800	100	100	100
4200	100	100	100
4600	100	100	100
5000	100	100	100

La figure 1-10 montre le pourcentage de simulations qui convergent pour chaque méthode d'échantillonnage. Pour la méthode Aléatoire simple (SRS) et l'allocation proportionnelle dans la méthode stratifiée (PSRS), les pourcentages concernés sont à peu près égaux. Toutefois, pour l'allocation de Neyman (optimum) dans la méthode stratifiée (OSRS), ce pourcentage est toujours inférieur à ceux des deux autres méthodes.

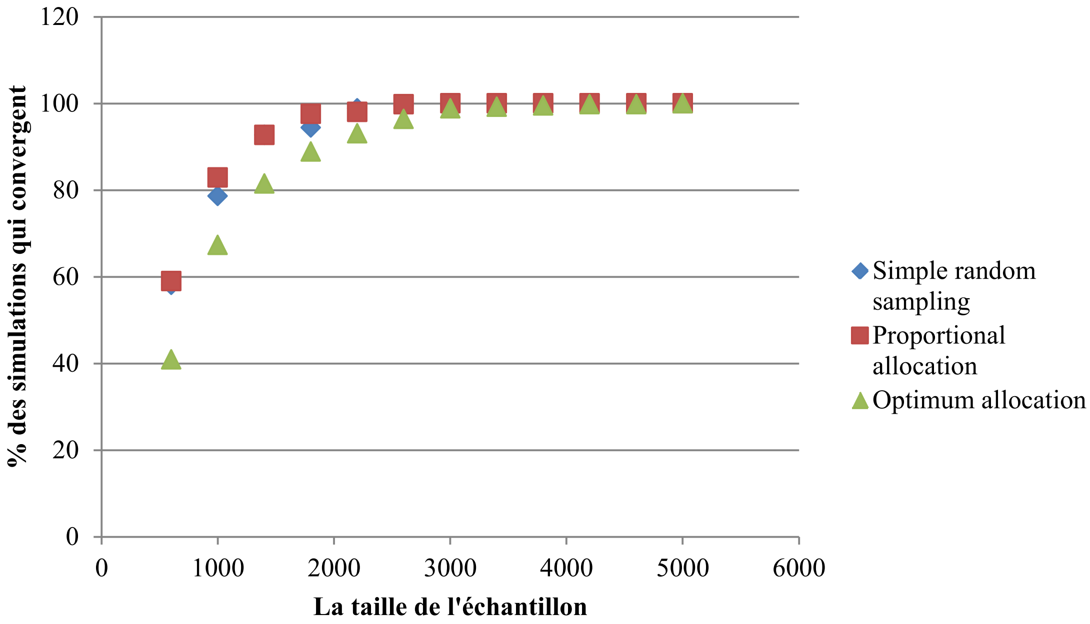


Figure 1-10: Pourcentage des simulations qui convergent vs. la taille de l'échantillon

Les figures 1-11, 1-12 et 1-13 montrent respectivement l'exactitude (*accuracy*), la précision et la signification statistique des coefficients d'après les simulations numériques. Ces statistiques s'améliorent lorsque la taille de l'échantillon augmente. Toutefois, il est très important de noter qu'avoir un échantillon représentatif ne signifie pas nécessairement que le modèle multi-variable converge avec succès.

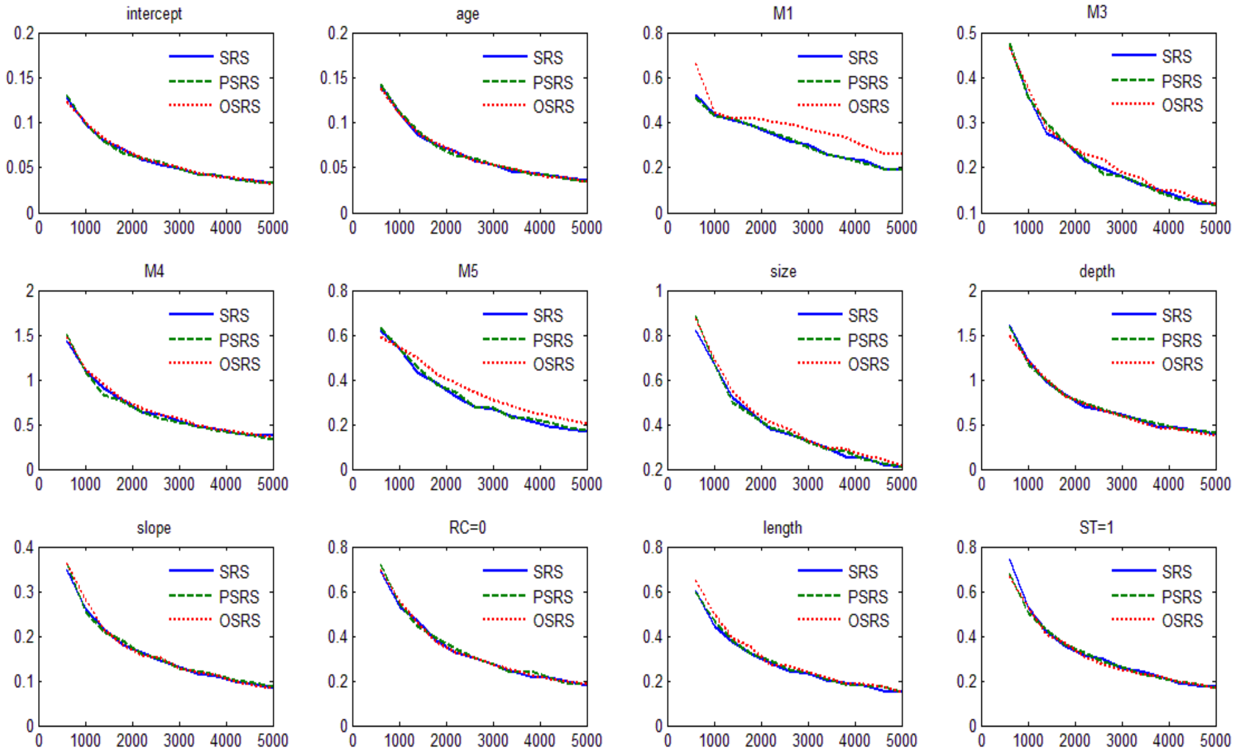


Figure 1-11: Erreur relative moyenne (*average relative bias*) des coefficients estimés vs. taille de l'échantillon (pour un échantillon : l'erreur est calculée en comparant le coefficient estimé sur cet échantillon avec le coefficient estimé sur la population totale) pour les différents facteurs explicatifs présumés, dont M1: Brick, M2: vitrified clay, M3: clay tile, M4: béton armé, M5: béton, RC : type de rue, ST : type de réseau (1 : unitaire).

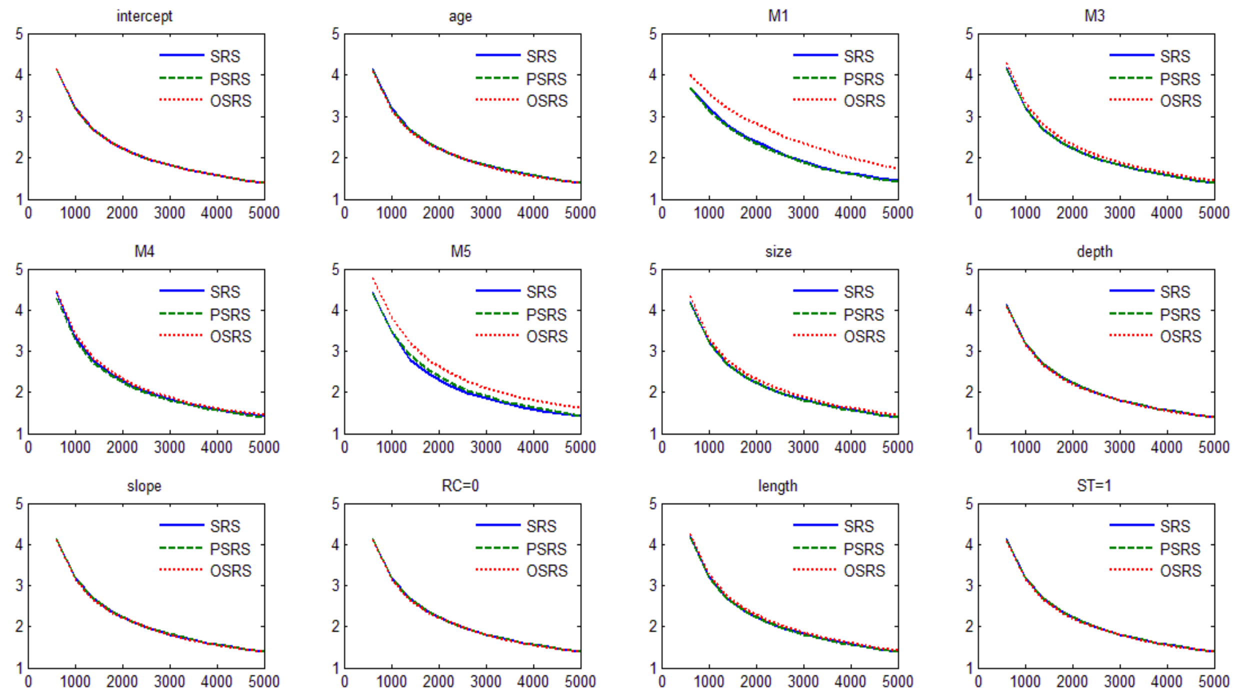


Figure 1-12: moyenne des erreurs types des coefficients estimés vs. taille de l'échantillon, pour les différents facteurs explicatifs présumés, dont : M1: Brick, M2: vitrified clay, M3: clay tile, M4: béton armé, M5: béton, RC : type de rue, ST : type de réseau (1 : unitaire).

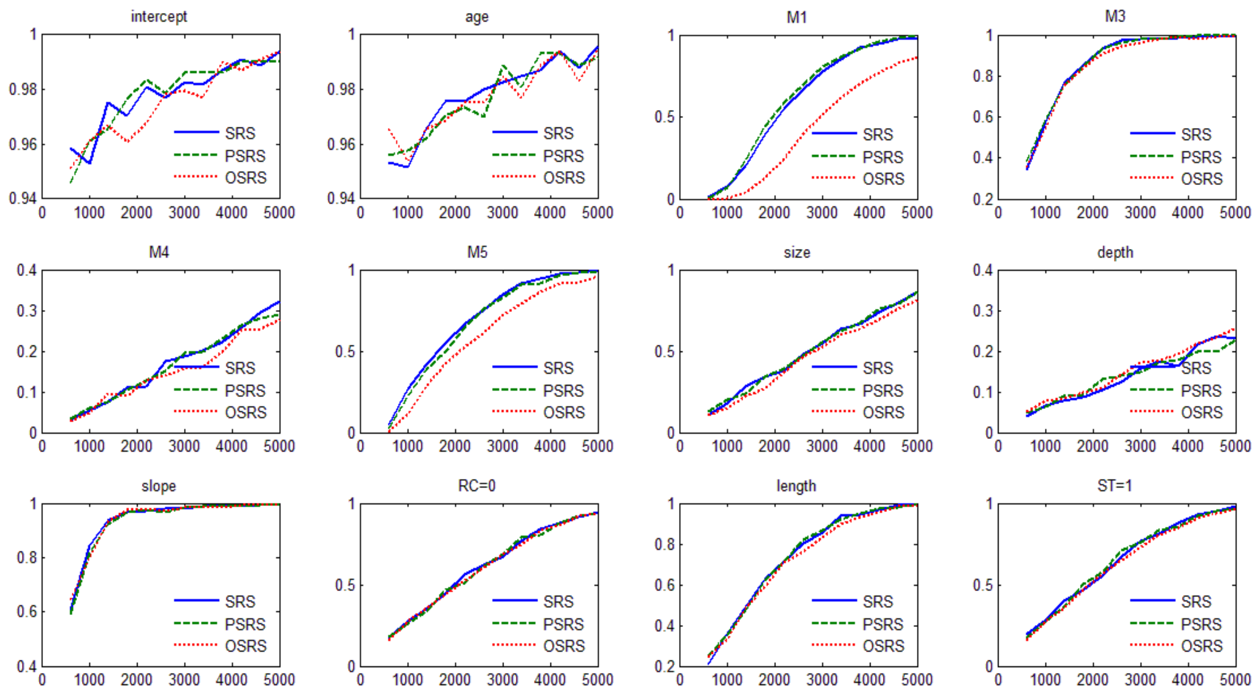


Figure 1-13: Proportion des simulations pour lesquelles (pour un niveau de confiance de 90%) l'intervale de confiance autour de la valeur estimée : a) contient la valeur estimée sur la population mère et b) ne contient pas la valeur zero ; étude pour les différents facteurs explicatifs présumés, dont M1: Brick, M2: vitrified clay, M3: clay tile, M4: béton armé, M5: béton, RC : type de rue, ST : type de réseau (1 : unitaire).

La qualité du calage dépend donc des points suivants :

- La méthode d'échantillonnage utilisée ;
- La taille de l'échantillon disponible (nombre de tronçons inspectés et aux caractéristiques connues) ;
- Le nombre de tronçons en mauvais état dans le patrimoine étudié (dans le cas d'un état de santé binaire) ;
- Le nombre de facteurs explicatifs présumés (nombre des variables disponibles) ;
- La qualité des données disponibles.

1.5. Perspectives de recherche

Les différents volets de cette recherche ont également conduit à identifier un ensemble de perspectives. Tout d'abord concernant l'élaboration des programmes d'inspection, il serait judicieux d'étudier l'influence de la taille du stock initial (de tronçons inspectés) sur la précision du modèle de détérioration obtenu et donc sur l'efficacité du programme d'inspection. Il serait également intéressant d'évaluer le bénéfice, en termes économiques, de chaque variante ; cela suppose de pouvoir estimer les coûts liés à l'acquisition de données et les bénéfices ou coûts évités par un ciblage efficace des tronçons défaillants.

Concernant l'évaluation de l'état de santé d'une conduite grâce au rapport d'inspection télévisée, il sera nécessaire d'enrichir la base de connaissance servant à calibrer les seuils (notamment en réalisant de nouvelles campagnes d'évaluation de tronçons par les experts).

Concernant l'évaluation d'un patrimoine à partir d'un échantillon, les premiers résultats obtenus sont prometteurs, il s'agit maintenant de pouvoir tester d'autres méthodes d'échantillonnage (comme par exemple l'échantillonnage en grappe). La question suivante sera de relier le choix de la méthode d'échantillonnage au patrimoine considéré ; par exemple, les réseaux de neurones artificiels ne sont pas adaptés à des petites bases de données (Trans *et al*., 2009).

1.6. Référence

1. Agresti, A., 2002. "Categorical data analysis", 2nd Ed. Hoboken, NJ: Wiley.
2. Ahmadi, M., Cherqui, F., De Massiac, J.C., Le Gauffre P. (2013). "From sewer inspection programs to rehabilitation needs: research and results related to data quality and availability with the support of numerical experiment". *European Journal of Civil Engineering* (In press).
3. Baur, R. and Herz, R. (2002). "Selective inspection planning with ageing forecast for sewer types". *Journal of Water Science and Technology*. 46, 4-7, 389-396.
4. CEN, 2003, EN 13508-2: "Condition of drain and sewer systems outside buildings, Part 2: Visual inspection coding system", European Norms.
5. Gemora, I., (2003). PACP: "Pipeline Assessment and Certification Program". Proc. Int. Conf. on Pipeline Engineering and Construction, ASCE, Reston, VA, 822–825.
6. Le Gauffre, P., Cherqui, F., De Massiac, J.C., Joannis, C., Werey, C., Rozan, A. (2010). "Indicateurs de performance pour la gestion patrimoniale des réseaux d'assainissement urbains". Lyon, France.
7. Le Gauffre, P., Joannis, C., Vasconcelos, E., Breysse, D., Gibello, C. and Desmulliez, J. J. (2007). "Performance indicators and multicriteria decision support for sewer asset management". *Journal of Infrastructure Systems*, 13 (2), 105–114.
8. Le Gauffre, P., Joannis, C., Breysse, D., Gibello, C. and Desmulliez, J.J., (2004). "Gestion patrimoniale des réseaux d'assainissement urbains", Guide méthodologique. Paris, Lavoisier Tec&Doc, 416 p., ISBN 27430-0748-6, (in French).
9. Tran, D.H., Perera, B.J.C., and Ng, A.W.M., (2009). "Comparison of structural deterioration models for stormwater drainage pipes". *Computer-Aided Civil and Infrastructure Engineering*, 24 (2), 145–156.

9.1. Water Research Center (WRc). (2004). "Sewerage rehabilitation manual", 4th Ed., WRc, Wiltshire, U.K.

10. Wirahadikusumah, R., Abraham, D., Iseley, T. (2001). "Challenges in modeling deterioration of combined sewers." *Journal of Infrastructure Systems,*7(2) 77-84.

Chapter II: Introduction and overview

2.1. Broad context of the thesis

Sewer systems form one of the most capital-intensive infrastructure systems in the developed countries. According to a survey performed by Malik *et al*. (1997) for the U.S., an average city or sanitation district has 1,075 km (667 mi) of sewer, a population of 221,199 and an annual budget of almost $3 million. These figures are equivalent to a spending of $18 per person or $3,435/km of sewer (Wirahadikusumah *et al*., 2001). While the average age of sewers is 42 years (in 1997), almost half of the municipalities across the United States have sewers with an average age of more than 50 years (in 1997). To improve the condition of these old and deteriorating systems, the EPA (ASCE 1998) has estimated that $10 billion is needed for upgrading existing wastewater collection systems and $45 billion for controlling combined sewer overflows.

In France, according to Lesage (2013) as the sewer network which is about 360,000 km, is aging, vast rehabilitation actions are required by 2020. In total, the estimated annual cost may be about 3 billion Euros.

Hence, to safeguard proper functioning of the sewer systems over time, they need to be managed cost-effectively. Traditionally, municipalities have addressed the design, construction, maintenance, and operation of sewer systems with a crisis-based (reactive or run-to-failure) approach (Wirahadikusumah *et al*., 2001). This practice results in the inefficient use of limited funds, causing more frequent sewer failures that end in difficult and costly rehabilitations.

According to WERF (2009), finding the optimal solution to address problems in sewer systems has always challenged asset managers. An integrated approach for the determination of deterioration of sewers is necessary to fully gauge the condition of these underground systems. This involves (Saegrov 2006; WERF 2007):

- Programming inspection needs;
- Routine and systematic sewer structural and hydraulic condition assessments by different inspection techniques;
- Establishment of a standard condition rating system;
- Developing and updating prediction models for sewer condition;
- Rehabilitation techniques.

An integrated sewer system management (proactive approach), which includes life cycle cost analysis and development of prioritization schemes for selection of inspection and rehabilitation needs is essential for cost-effective asset management.

In this manuscript, we adapt a specific definition of a proactive sewer asset management from a utility point of view (figure 2-1 adapted from Ahmadi *et al*., 2013). Figure 2-1 illustrates main steps of this definition containing 4 different levels of asset management practices existing within utilities from a simple run-to-failure model to a long-term proactive asset management politics.

These levels are as follows:

(1) Defining rehabilitation needs by determination of inspected segment's condition;

(2) Prioritizing rehabilitation needs based on a risk approach by combining information about observed dysfunctions and network's environment variables with assessed condition grade (These two levels are limited to segments already inspected);

(3) Prioritizing inspection needs by considering one or the combination of the followings:

(a) Vulnerability of segments' environment;

(b) Estimated condition of sewer segments;

(4) Long-term planning of inspection/ rehabilitation and budget needs from a representative sample of the asset stock.

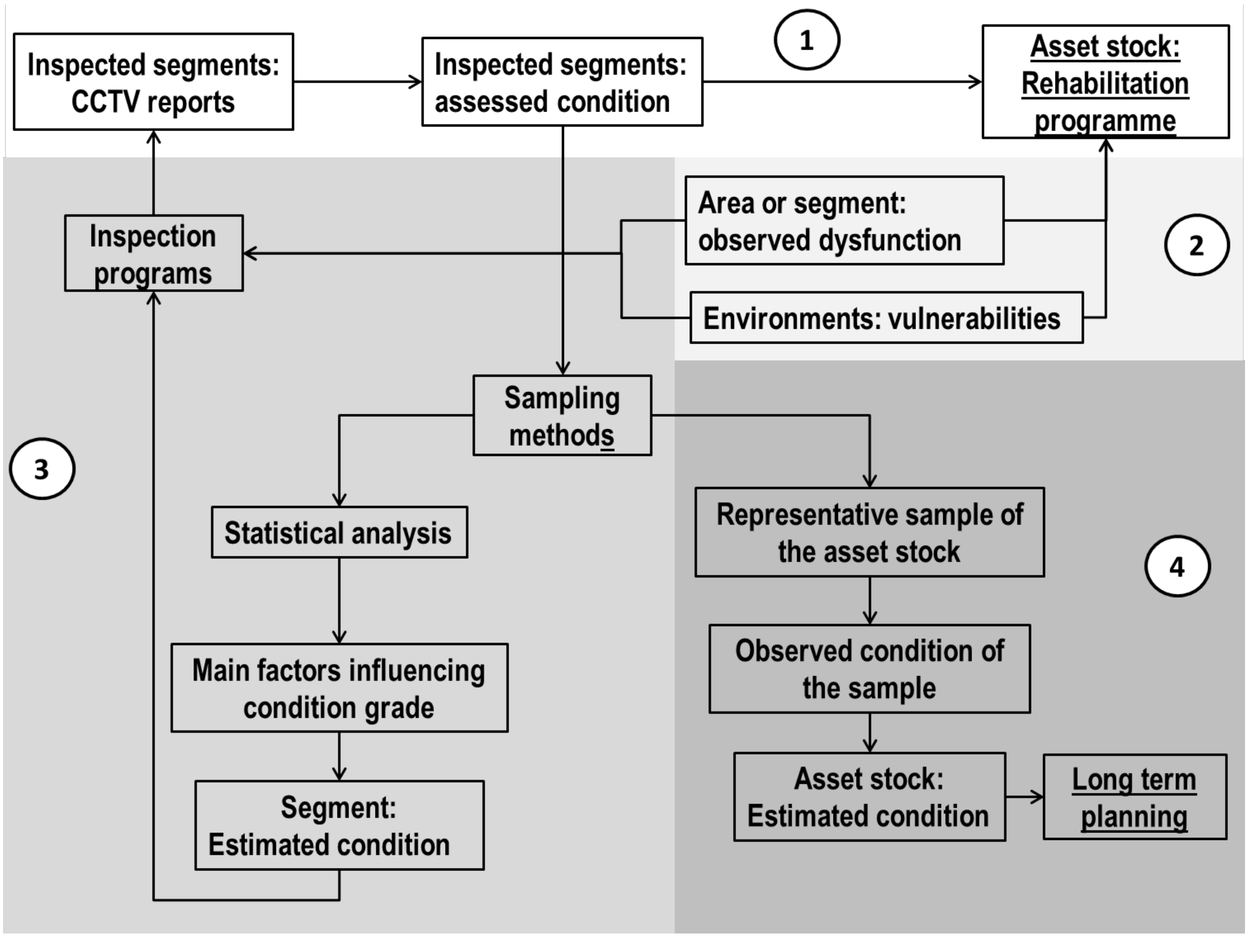


Figure 2-1. Sewer asset management (SAM) levels of practices, the simplest practice is in light color (1) and each different grey color corresponds to another improvement of the practices (from the lightest to the darkest).

The first and second levels were developed as parts of French projects RERAU and INDIGAU (cf. 2.2) based on CCTV inspections according to EN15308-2. These works as well as the definition of certain dysfunction indicators along with the provision of three distinct toolboxes dedicated to the use of utility managers were carried out between 2001 and 2010 (Le Gauffre *et al*., 2004; Le Gauffre *et al*., 2007; Le Gauffre *et al*., 2010).

As it is mentioned above, these levels of sophistication about asset management, however, are limited to inspected segments. In other words, all uninspected segments are excluded from the decision-making process of rehabilitation works. On the other hand, normally the inspection surveys are performed based on either the judgment of the CCTV operators within the areas susceptible of problems such as odors, infiltration etc. or reasons with a random-nature such as road works etc. Hence, the need of elaborating inspection programs more cost-effectively is felt.

However, few approaches are available as decision-making tools for planning inspection surveys (Baur and Herz 2002). High numbers of deficient segments resulting from poor maintenance activities and budget restrictions necessitate the development of prioritization tools to address inspection needs of segments with the highest risk of failure (Salman 2010).

In order to assess a segment's risk of failure, its probability of failure should be combined by its consequences of failure (Le Gauffre *et al.*, 2007). The probability of failure could be assessed by deterioration models. Hence, these inspection programs are based generally on deterioration models allowing the prediction of current and future condition state of assets from the factors known having an influence on the condition of the assets (c.f. section 3.2.4.1)..

Various types of deterioration models could be found within scientific literature (c.f. section 3.2.4.2.). However, despite the development of various deterioration models, attention is still focused on the type of deterioration model used and the influence of available data on the predictive power of these models in terms of data uncertainty, imprecision and incompleteness has gone unstudied. Main research questions associated with quality of data which were dealt with within this thesis are then as follows:

(1) Which influential factors are most informative? Could we establish the list of most informative variables in terms of inspection programs?

(2) Is it worth to accept a degree of imprecision instead of incompleteness within the utility database?

(3) How can the data most probably available within a utility be used to define an effective inspection program? For example, are data collected for hydraulic models of network such as segment's diameter, depth, gradient, sewer type and length, sufficient and relevant enough for inspection programs?

(4) Can we use an auxiliary variable in order to compensate effects of missing data?

(5) Is it worth to accept a degree of uncertainty within data instead of not having them?

In addition, once segments are inspected, they should be evaluated by a condition grading protocol which should remain identic for all segments. Though various condition grading protocols have been developed, they all fail to address the following point which is addressed in this manuscript:

- The condition assessment process contains some uncertainty sources. The condition assessment process for a segment could result in an over-estimation or an under-estimation of the real condition grade of the segment, However, the existing condition grading protocols do not allow the stakeholders to take into account their own specifications about the under-estimation or over-estimation of their assets' conditions. Therefore, the need of a more specific condition grading protocol is felt which could take into account the sensitivity of utility managers and stakeholders to this specific issue.

In fact, at the moment small number of utilities has completely inspected and evaluated their asset stocks. Therefore, the use of a representative sample from an asset stock in order to calibrate decision-support models as deterioration models, to study scenarios about future and to draw appropriate conclusions for the whole asset stock seems mandatory. Nevertheless, in this regard we should tackle with following problematic issues:

- How to draw a representative sample of an asset stock which reflects in the best and appropriate manners the characteristics of this asset stock?
- How to provide a reliable estimation of a specific property of the asset stock from this sample?
- Calibration of multivariate models seems somehow tricky. Therefore, what is the impact of used sample on the calibration outcomes of these multivariate models?

2.2. Theoretical and operational basis of the thesis

As it is mentioned above, this thesis is the continuation of French research projects called RERAU and INDIGAU. The following paragraphs describe briefly the main outcomes of these projects on which this manuscript is based.

2.1.1. RERAU

During the last ten years an increasing number of studies has been dedicated to water and sewer asset management. At an international level, working groups of the International Water Association (IWA) proposed performance indicators for water (Alegre *et al*., 2000) and for sewer systems (Matos *et al*., 2003).

In North America, the works of NRC-CNRC and INRS in Canada have been particularly important (Vanier *et al*., 2004, 2006, 2006b; Mailhot *et al*., 2000; Rahman and Vanier, 2004). Several tools have also been developed and supported by WERF – *Water Environment Research Foundation*, such as SCRAPS (Sewer Cataloging, Retrieval, and Prioritization System) dedicated to the definition of inspection needs and priorities, for sewer pipes (Hahn *et al*., 2002; Merrill *et al*., 2004; Marlow *et al*.,2007).

In Europe, several large R&D projects were supported by the European Union, including:

- CARE-W : Computer-Aided Rehabilitation of Water networks (Saegrov, 2005) ;
- CARE-S : Computer Aided Rehabilitation for Sewer and Stormwater Networks (Saegrov, 2006) ;
- APUSS : Assessing Infiltration and Exfiltration on the Performance of Urban Sewer Systems (Ellis & Bertrand-Krajewski, 2006).

In the UK, the "whole life costing" approach was implemented through software tools for the UK context (Cashman *et al*., 2004).

In France, on part of the national RERAU program ("Réhabilitation des Réseaux d'Assainissement Urbains") was devoted to the description and construction of a set of criteria for the rehabilitation and inspection of sewerage systems. The proposed methodology is detailed in Le Gauffre *et al*., (2004) and is presented in Le Gauffre *et al*., (2007). This contribution to asset management of sewer segments may be summarized as follows:

- Two sets of criteria are defined for supporting the definition of investigation and rehabilitation programs;
- Criteria are calculated from complementary performance indicators (PIs) relating to defects, dysfunctions or impacts (table 2-1);
- PIs may be assessed by observation, by estimation, or by combination of other PIs;
- PIs refer to individual assets or to subsystems;
- Each PI is assessed on an ordinal four-grade scale. Also, relationships between PIs are modeled by qualitative models (4x4 tables);
- Each proposed criterion assesses a contribution of a particular dysfunction of a sewer segment to a particular impact (table 2-1). Each of the 8 defined impacts is linked to some of the 10 source dysfunctions (figure 2-1).

Decision criteria are related to a three-level causal chain linking defects to their impacts (Le Gauffre *et al*., 2007). Existing defects in a segment characterize the deviation of actual physical condition from the standard condition of an asset. These defects lead to consequences for facility operation, known as dysfunctions (table 2-1). To allow such raw information to be more easily used, all indicators are measured on an ordinal scale of increasing severity from G1 to G4.

At the end, these dysfunctions may result, according to the asset's context, to an impact. The context components taken into account to evaluate impacts are known as "vulnerability factors". Le Gauffre *et al*., (2004) shows how to construct 30 investigation and 31 rehabilitation criteria by using 64 dysfunction indicators plus 12 vulnerability indicators.

Table 2-1. Dysfunction and impact indicators assessed from visual inspection reports

Indicator of dysfunction	**Definition**	**Indicator of impact**	**Definition**
ABR	Ongoing degradation from abrasion	DAB	Damage to buildings (including infiltration of water into basements)
BLO	Blockage	NUH	Nuisance of a hydraulic nature (service interruption, flooding…)
COL	Risk of collapse	OCP	Treatment plant operating surplus costs
COR	Degradation due to corrosion	OCS	Network operation surplus costs (including the cost of equipment shortened lifetime)
CSO	Excessive spillage	POG	Pollution of ground and groundwater
EXF	Exfiltration (seepage loss)	POL	Pollution of surface water resources, due to overflows, spillage or disturbance to treatment system process
FLO	Flooding	SLC	Shortened lifetime cost, along with the surplus cost associated with remedial actions
HYD	Decrease in hydraulic capacity	TRA	Traffic disruption (including for the needs of sewer operation: Cleansing …)
INF	Infiltration		
ROO	Ongoing degradation from root intrusion		
SAN	Sand silting		
SPD	Destabilization of ground-pipe system		

Section 5 of European standard EN 752, providing a framework relating to planning, construction, rehabilitation, maintenance and operation for drain and sewer systems outside buildings, defines the functional requirements of sewer networks. Table 2-2 lists all these requirements and gives also the related RERAU project dysfunctions and impacts. For example, the water tightness requirement can be taken into account with two dysfunctions (Infiltration and Exfiltration) and 5 impacts (Pollution of surface water resources, pollution of ground and groundwater, surplus cost of network operation, surplus cost of wastewater treatment plant operation and damage to buildings) in the RERAU methodology (Le Gauffre *et al*.,2004).

Table 2-2. EN 752 requirements and related RERAU project dysfunctions

Functional requirement of EN752 (2008)	**Related RERAU dysfunction**	**Related RERAU impacts**
Protection from flooding	HYD, BLO, FLO	NUH
Maintainability	-	-
Protection of surface receiving waters	INF, HYD, BLO, CSO	POL, POG
Protection of groundwater	EXF	POL, POG
Prevention of odors and toxic, explosive and corrosive gases	-	NUH
Prevention of noise and vibration	-	-
Sustainable use of products and materials	-	-
Structural integrity and design life	SPD, ROO, COR, ABR	SLC
Maintaining the flow	BLO	POL, POG, NUH, OCS
Water tightness	INF, EXF	POL, POG, OCS, OCP, DAB
Not endangering adjacent structures and utility services	COL	DAB, TRA
Inputs quality	-	-

R/POL/BLO, R/OCP/INF, R/TRA/COL and R/SLC/SPD are four examples of decision criteria that may be used to define rehabilitation needs and priorities (figure 2-2):

- R/POL/BLO: sewer segment contributing to pollution of surface water (POL) due to spillages induced by repeated blockages (BLO);
- R/OCP/INF: sewer segment contributing to infiltration (INF) inducing treatment plant operation surplus costs (OCP);
- R/TRA/COL: risk of traffic disruption (TRA) due to collapse (COL);
- R/SLC/SPD: shortened lifetime cost (SLC) due to destabilization of the ground-pipe system (SPD).

DYSFUNCTIONS

IMPACTS	INF	EXF	HYD	SAN	BLO	SPD	COR	ROO	ABR	COL
POL	X		X	X	X					
POG		X	X	X	X					
NUH			X	X	X					
TRA				X	X					X
DAB		X								X
OCS	X			X	X					
OCP	X									
SLC						X	X	X	X	

Figure 2-2. Decision criteria defined by linking impacts and source dysfunctions

2.1.2. INDIGAU

The INDIGAU R&D project (2007-2010), financed by the French National Research Agency, involved four research centers, one private company (*G2C environnement*) and seven urban utilities (Werey *et al*., 2010). The objective was to produce a collaborative decision-support tool to perform automated condition assessment, multi-criteria sorting of sewer segments, and to determine rehabilitation programs from the methodology developed during the RERAU national French project. This toolbox intended for the use of asset managers and stakeholders is available on a web interface providing three functionalities as follows:

1. *Data conversion*: The INDIGAU toolbox introduces a preliminary tool for utilities that have CCTV data coded in former coding systems. It translates the data into EN 13508-2 codes readable in the other INDIGAU tools. Indeed, normalized codes started to be widely used after 2008 even though the European standard has been released in 2003. The conversion tool is semi-automatic and self-learning. By introducing a specific dysfunction, this tool saves the name, description and other introduced details for next uses.

2. INDIGAU performs automated interpretation of CCTV inspections to calculate sewer dysfunction indicators, using interpretation models, on a four-grade scale. The calibration of the interpretation model against an expert-opinion database removes the risk of single-human bias (Cherqui *et al,*. 2008 and chapter 5).
3. The INDIGAU toolbox offers a criteria construction workshop and full control of parameters for indicators combination and multi-criteria analysis for prioritizing the segments to be rehabilitated. The outcome of the process is a sorting of individual sewer segments into three classes corresponding to priority of rehabilitation.

2.3. Thesis outlines

The remainder of this manuscript is organized as follows. In chapter 3, we will explain asset management definitions and proactive programs in water and wastewater utilities and municipalities. We will also re-explain in-detail our point of view of proactive approaches in sewer asset management. Afterwards, the need of inspection programs will be discussed. Then, the assessment of a segment into a condition grade according to different existing assessment methods will be unfolded. The final part is dedicated to the need of assessing an asset stock as a whole and to study the impact of our in-time knowledge on the decision-support models of sewer asset management.

In Chapter 4, we will tackle problematic issues about inspection programs as well as data quality on their efficiency.

Chapter 5 describes our proposed condition grading protocol which could take into account the sensitivity of utility managers and stakeholders to over or underestimation of a segment's condition grade.

Chapter 6 focuses on the assessment of an asset stock from a sample and consequences that could have the use of a representative sample of an asset stock.

Finally chapter 7 will summarize the main conclusions and opens some windows for further research.

2.4. References

11. Alegre, H., Hirner, W., Baptista, J.M., & Parena, R. (2000). “Performance indicators for water supply services IWA (International Water Association) manual of best practice”. London: IWA Publishing.
12. American Society of Civil Engineers. (1998). ‘‘Report Card for America’s Infrastructure.’’ ASCE Annual Report, Reston, Va.
13. Ahmadi, M., Cherqui, F., De Massiac, J.C., Le Gauffre P. (2013). “From sewer inspection programs to rehabilitation needs: research and results related to data quality and availability with the support of numerical experiment”. *European Journal of Civil Engineering* (In press).
14. Baur, R. and Herz, R. (2002). “Selective inspection planning with ageing forecast for sewer types”. *Journal of Water Science and Technology*. 46, 4-7, 389-396.
15. Cashman, A., Savic, D., Saul, A., Ashley, R., Walters, G., Blanksby, J., Djordjevic, S., Unwin, D., Kapelan, Z., (2004). “Whole life costing of sewers (Cost-S): developing a sustainable approach to sewer network management”, Proceedings Novatech, Lyon, France.
16. CEN, 2003, EN 13508-2: “Condition of drain and sewer systems outside buildings, Part 2: Visual inspection coding system”, European Norms.
17. Cherqui, F., Werey, C., Ibrahim, M., Le Gauffre, P., (2008). “CCTV inspection of sewer segments: calibration of performance indicators based on experts’ opinions”. Proceedings of the 11th International Conference on Urban Drainage, Edinburgh, UK, 31 Aug.-5 Sept., 10 p.
18. Ellis, J.B. and Bertrand-Krajewski, J.-L. (Ed.), (2006). "Assessing infiltration and exfiltration on the Performance of Urban Sewer Systems (APUSS)". IWA Publishing, London, 192 p.
19. EN 752: 2008, *Drain and sewer systems outside buildings.* European Norms.
20. Hahn, M., Palmer, R., Merrill, M., Andrew, B., Lukas, A., (2002). “Expert system for prioritizing the inspection of sewers: Knowledge base formulation and evaluation”. ASCE, *Journal of Water Resources Planning and Management*, 128(2), 121-129.
21. Le Gauffre, P., Cherqui, F., De Massiac, J.C., Joannis, C., Werey, C., Rozan, A. (2010). “Indicateurs de performance pour la gestion patrimoniale des réseaux d’assainissement urbains”. Lyon, France.
22. Le Gauffre, P., Joannis, C., Vasconcelos, E., Breysse, D., Gibello, C. and Desmulliez, J. J. (2007). “Performance indicators and multicriteria decision support for sewer asset management”. *Journal of Infrastructure Systems*, 13 (2), 105–114.
23. Le Gauffre, P., Joannis, C., Breysse, D., Gibello, C. and Desmulliez, J.J., (2004). “Gestion patrimoniale des réseaux d’assainissement urbains”, Guide méthodologique. Paris, Lavoisier Tec&Doc, 416 p., ISBN 27430-0748-6, (in French).

24. Lesage, M. (2013). “Rapport d’évaluation de la politique de l’eau en France: Mobiliser les territoires pour inventer le nouveau service public de l’eau et atteindre nos objectifs de qualité”. Ministry of Homeland, Paris, France (In French).
25. Mailhot, A., Duchesne, S., Musso, E., Villeneuve, J.P., (2000). “Modélisation de l’évolution de l’état structural des réseaux d’égout : application à une municipalité du Québec”. *Canadian Journal of Civil Engineering*. 7 , 65-72.
26. Malik, O., Pumphrey, N. D., and Roberts, F. L. (1997). ‘‘Sanitary sewers: State-of-the-practice.’’ *Proc., Conf. on Infrastruct. Condition Assessment:Art, Sci., and Pract.*, ASCE, New York, 297–306.
27. Marlow, D., Davis, P., Trans, D., Beale, D., Burn, S., Urquhart, A., (2009). “Remaining asset life: a state of the art review”, WERF – Water Environment Research Foundation, IWA Publishing, London, UK.
28. Marlow, D., Heart, S., Burn, S., Urquhart, A., Gould, S., Anderson, M., Cook, S., Ambrose, M., Madin, B., & Fitzgerald, A. (2007). “Condition assessment strategies and protocols for water and wastewater utility assets”, WERF – Water Environment Research Foundation, London: IWA Publishing, p. 476.
29. Matos, R., Ashley, R.M., Cardoso, A., Molinari, A., Schulz, A., & Duarte, P. (2003). “Performance indicators for wastewater services”. London: International Water Association, ISBN: 1900222906.
30. Merrill, S., Lukas, A., Palmer, R.N., Hahn, M.A., (2004). “Development of a tool to prioritize sewer inspections”. Technical report, WERF - Water Environment Research Foundation, 123 p.
31. Rahman, S. and Vanier, D.J., (2004). “MIIP Client Report: An Evaluation of Condition Assessment Protocols for Sewer Management”, National Research Council of Canada, Ottawa.
32. Sægrov,S., eds. (2006). “Care-S Computer Aided Rehabilitation of Sewer and Storm Water Networks.” IWA Publishing, London, UK.
33. Saegrov, S. (Ed.), 2005. “CARE-W: Computer Aided Rehabilitation for Water Networks”, IWA Publishing, London, 208 p.
34. Salman, B., (2010). “Infrastructure management and deterioration risk assessment of wastewater collection systems”. Thesis (PhD), Univ. of Cincinnati, OH.
35. Vanier, D.J, Newton, L.A., Rahman, S., (2006). “A Framework for municipal infrastructure management for Canadian municipalities”. Building Science Insight conference.
36. Vanier, D.J. and Newton, L.A., (2006b). “MIIP Report: The State of Canadian Sewers - Analysis of Asset Inventory and Condition. Client Report”. NRC Institute for Research in Construction, 38 p.
37. Vanier, D.J. and Rahman, S., (2004). “MIIP Report: A Primer on Municipal Infrastructure Asset Management. Client Report”. NRC Institute for Research in Construction, 71 p.
38. Wirahadikusumah, R., Abraham, D., Iseley, T. (2001). “Challenges in modeling deterioration of combined sewers.” *Journal of Infrastructure Systems,*7(2) 77-84.

Chapter III: Literature review

The future-retrofitting benefits of asset management approaches have attracted attention by researchers, government agencies and public and private sector managers (Syachrani 2004). Many researches and publications are hence devoted to the development of asset management regulations, policies, practices and lessons learned from the applications of these approaches in different cities or countries (WERF 2007).

These researches have been carried out in order to study and explain the process of deterioration, condition grading, identification of structural, operational and hydraulic factors influencing the condition grade of an asset and prediction of future condition of assets. This chapter explains the asset management and proactive programs in water and wastewater utilities and municipalities. We will also explain our point of view of proactive approaches in sewer asset management. Afterwards, the need of inspection programs will be discussed. Then, the assessment of a segment into a condition grade according to different existing assessment methods will be unfolded. The final part is dedicated to the need of assessing an asset stock as a whole and to study the impact of our in-time knowledge on the decision-support models of sewer asset management.

3.1. Asset management

3.1.1. Definitions and levels of sophistication

Asset management in the water and wastewater industry has been adapted from many successful implementations in other industries such as transportation and oil infrastructure management (Syachrani 2004, WERF 2007).

Burns *et al*. (1999) underline some main features of infrastructure assets: *"infrastructure assets (...) are defined functionally as assets that are not replaced as a whole but rather are* *renewed piecemeal by the replacement of individual components whilst maintaining the function of the system as a whole. Infrastructure assets have indefinite lives. Economic lives, however, can be assigned to individual components of an infrastructure system".* However, Schulting & Alegre (2007) add that infrastructure systems have to meet performance objectives that are evolving: *"the aim of the rehabilitation actions is not to recover the initial characteristics, but rather to provide the necessary characteristics so that the infrastructure performs according to the actual needs and expectations".*

According to WERF (2007), asset management remains an ill-defined term as many definitions exist in the literature. Following complementary definitions provide a description that contains all of available aspect within asset management approaches:

- *"Asset management in the water sector can be described as managing infrastructure capital assets to minimize the total cost of owning and operating them, while delivering the service levels customers desire"* (Schulting & Alegre, 2007);
- *"Systematic and coordinated activities and practices through which an organization optimally manages its physical assets and their associated performance, risks and expenditures over their life cycles for the purpose of achieving its organizational strategic plan"* (PAS 55);
- *"Asset management is a continuous process that guides the acquisition, use and disposal of infrastructure assets to optimize service delivery and minimize costs over the asset's entire life"* (U.S. EPA 2002);
- The IIMM manual (IPWEA, 2006), developed in Australia & New Zealand, introduces the notion of "Advanced Asset Management" and defines three management levels (Marlow & Burns 2008):
  - *"Asset management is the combination of management, financial, economic, engineering and other practices applied to physical assets with the objective of providing the required levels of service in the most cost-effective manner"*;
  - *"Advanced asset management employs predictive modeling, risk management and optimized decision-making techniques to establish asset lifecycle treatment options and related long term cash flow predictions"*;

- *"Strategic asset management (SAM) is concerned with setting overall policy, strategy, and budgets using broad estimation tools (...); tactical asset management (TAM) is concerned with setting priorities (...) and it is applied at a finer resolution than SAM and more detailed data are required"*; (figure 3-1).

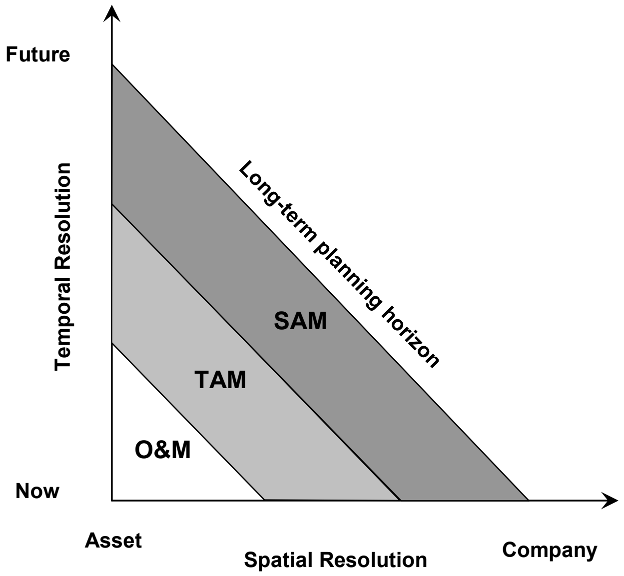


Figure 3-1. Different temporal and spatial scales of asset management, according to (Marlow & Burns, 2008)

3.1.2. Approaches to sewer asset management

Asset management approaches applied in the water and wastewater sectors of countries such as United States, Australia and United Kingdom can be characterized in terms of a succession of dominant philosophies (WERF 2007). The followings indicate the staged development of increasing asset management sophistication:

- Condition-based asset management: expenditure is focused on maintaining "what assets are and their conditions". If the condition is poor, the asset needs maintenance and/or rehabilitation to rectify defects.
- Performance-based asset management: "what assets do" is the main question in this level of sophistication. In other words, "is the asset doing the job that it was intended to do?" is the main question which can be often related to the asset's condition or may not.
- Service-based asset management: "is the asset contributing appropriately to the delivery of service" is the main question. Hence, this level of sophistication seeks to maintain the service provided by the asset stock at both local and regional level.

- Risk-based asset management: this level seeks to achieve optimum life cycle management of assets through consideration of risk to service provision with risk generally being defined as the product of 'probability of failure' and 'consequence of failure'.

Tactical and strategic asset management, both rely mostly on the assessment of inspection and rehabilitation priorities in mid and long-term programs (Rahman and Vanier 2004). These assessments involve the collection of data using inspection tools/techniques but other sources of data are also required to allow interpretation and contextualization of the results by considering the desired level of sophistication mentioned above.

Vanier (2001) notes that an effective asset management strategy in terms of asset maintenance in mid and long-term programs (tactical and strategic levels) generally consists of:

- Inspections that are carried out periodically to monitor and record how systems are performing;
- Preventive actions ensuring that systems or components will continue to perform their intended functions throughout their service life;
- Rehabilitation actions in order to replace one component of a system when it fails at the end of its service life;
- Capital renewal that replaces a system because of economic, obsolescence, modernization or computability issues.

Considering this definition, another classification could be defined for asset management approaches. The U.S. EPA (2002) identifies two different approaches:

- The advanced level asset management model or proactive approaches: components of assets are regularly maintained and finally replaced when deterioration outweighs;
- The run-to-failure management model or reactive approaches: assets are not regularly maintained. This led to higher replacement and emergency response costs.

Proactive approaches are, by definition, designed and carried out to (WERF 2007):

- Prevent failures before they occur

- Detect the onset of failures before they have an impact on the performance of the system

In this work, we adapt a specific definition of a proactive sewer asset management from a utility point of view (Ahmadi *et al*. 2013). Figure 3-2 illustrates main steps of this definition. This figure shows 4 different levels of asset management practices existing within utilities from a simple run-to-failure model to a long-term proactive asset management politics.

The simplest possible practice (process 1) occurring in a utility is to inspect a portion of the asset stock, then to determine their condition grade and finally to rehabilitate the inspected sewers in worst condition grade according to budget or soft-skill constraints available within a utility (run-to-failure approach).

In the first evolution (process 2), observed dysfunctions and network's environment variables are combined with assessed condition grade in order to prioritize the rehabilitation need for inspected sewers responsible for an impact. An impact reflects the degree to which dysfunctions induce noxious effects, depending on the context (Le Gauffre *et al*., 2007). By definition, within process 1&2, we are limited to the sewer segments already inspected. A sewer segment corresponds to a length of several meters which is homogeneous in some characteristics (material, diameter, etc.) and is mainly delimited by two successive manholes (Le Gauffre *et al*., 2007). Hence, in the next evolution, we should focus on inspections.

In the next evolution (process 3), prioritization of inspections can be based either on estimated condition of sewer segments and/or on environment's vulnerabilities in which they are located (segments located near or within vulnerable areas may be inspected first because dysfunction may lead to important impact). Estimating condition grade of segments requires the identification of factors influencing segments' conditions from a sample of already inspected segments and the use of prediction models. However, clearly these two could be combined.

The last evolution (process 4) contains the long term planning which is based on a representative sample of the asset stock. This latter allows the estimation of the whole asset stock's condition grade and consequently, to elaborate the best long term strategy (budget, inspection and rehabilitation needs).

As it is explained in chapter 2, next chapter is devoted to the definition of inspection programs and influence of data on their efficiency which is directly linked to the process (3) of figure 3-2. Chapter 5 deals with the problem of assessing a segment by an inspection directly linked to all processes and particularly to process 1. Finally, chapter 6 tackles the problematic issues of assessing an asset stock from a representative sample and calibration of decision-support models by this sample (process 3 and 4).

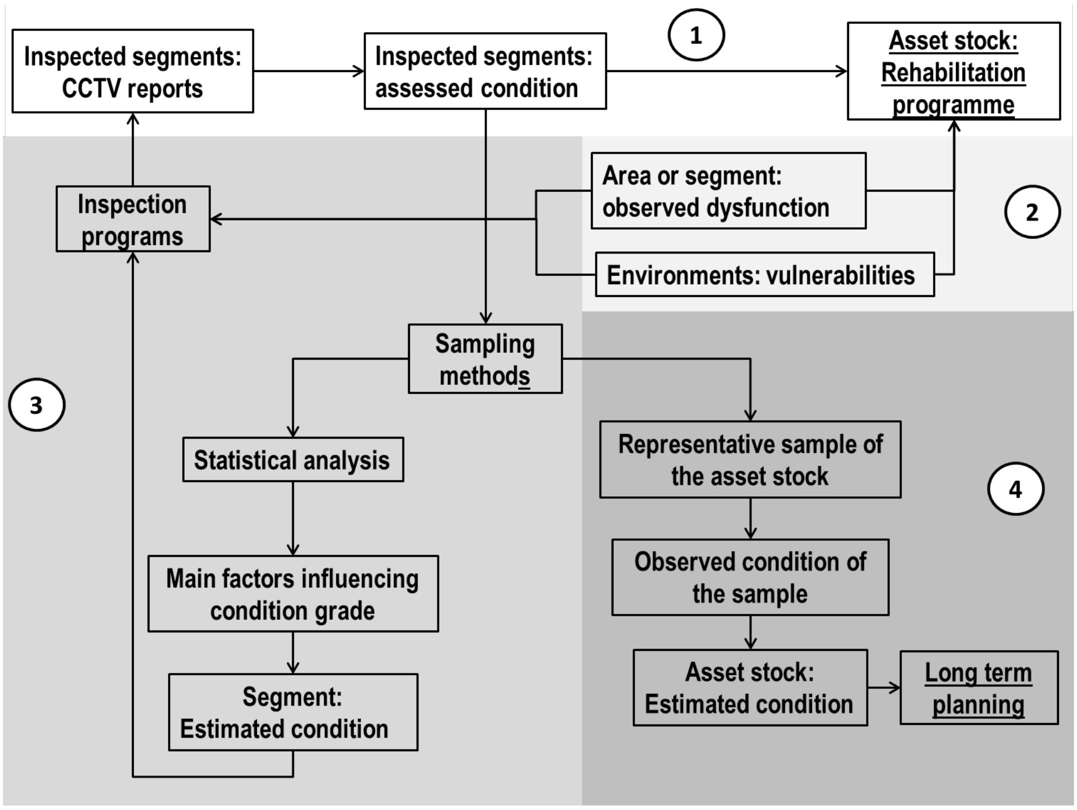


Figure 3-2. Sewer asset management (SAM) levels of practices, the simplest practice is in light color (1) and each different grey color corresponds to another improvement of the practices (from the lightest to the darkest).

Therefore, in this chapter, we will see first some aspects of inspection surveys containing existing inspection programs, inspection techniques, available data's role in these surveys and interest of using deterioration models in order to construct inspection surveys and their types. We will also describe certain available condition grading protocols and finally the need of assessing an asset stock from a sample is discussed.

3.2. Inspection programs

A vital component of these proactive approaches is the condition assessment of sewer segments (Ana and Bauwens 2010). Inspection programs should be established in order to detect and evaluate deterioration of assets due to in-service operation (WERF 2009). Development of an effective inspection program is centered on knowing when, where and how to inspect.

According to American Bureau of Shipping or ABS (2004) asset condition deteriorates over time and the level of deterioration gradually progress to the point that it can be detected (point *P* in figure 3-3). As time passes, the assets' conditions worsen and at the end they reach the failure point (point *F* in figure 3-3). Figure 3-3 is a hypothetical schematic curve of asset deterioration over time. However, it can vary in practice and in some cases be very inconsistent (WERF 2009).

Traditionally, inspection frequencies were set in terms of time-based or calendar-based intervals (Berardi *et al*., 2009). However; these traditional approaches do not consider asset's risk of failure on the required inspection interval. Consequently, the cost of inspection is not considered along with asset-related risk of failure.

However, few approaches are available as decision-making tools for planning inspection surveys (Baur and Herz 2002; Hahn *et al.,* 2002; Marlow *et al.,* 2007, 2008; Merrill *et al.,* 2004). High numbers of deficient segments resulting from poor maintenance activities and budget restrictions necessitate the development of prioritization tools to address inspection needs of the segments with the highest risk of failure (Salman 2010).

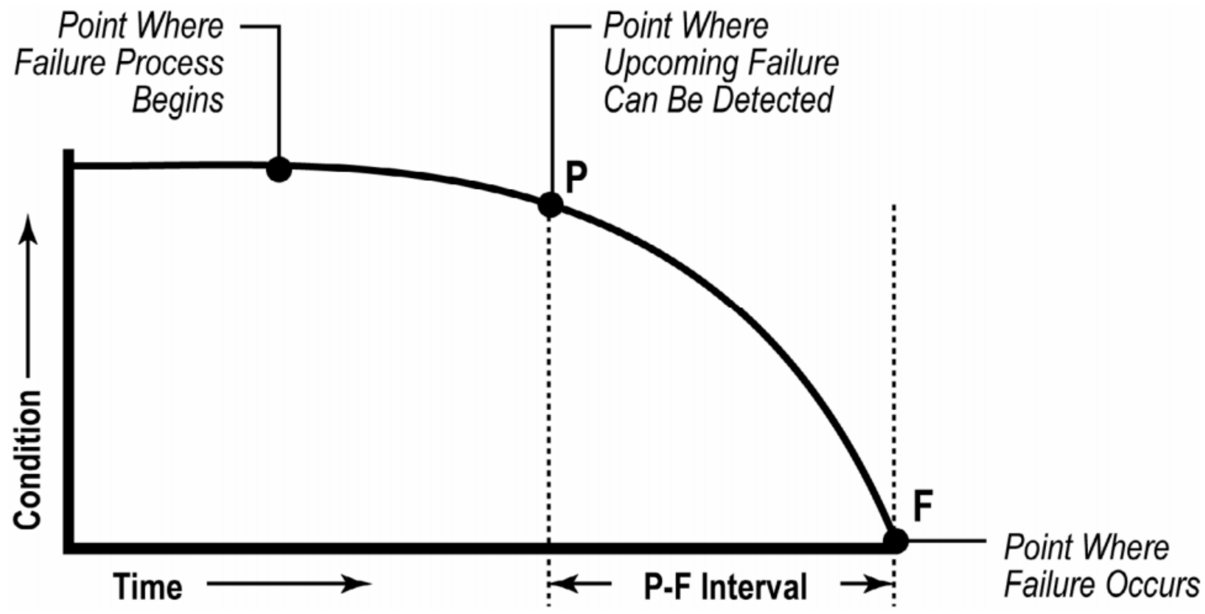


Figure 3-3. Failure process according to ABS (2004)

3.2.1. Risk-based approaches for inspection programs

According to the American Bureau of Shipping (ABS 2004), the main steps of the development of a risk-based approach are as follows:

- The determination of the risk of the potential failures of each asset
- The identification of the degradation mechanisms leading to component failures
- The selection of most effective inspection techniques allowing the detection of the progression of degradation mechanisms
- The development of an optimized inspection plan using the knowledge gained in the three previous steps
- The analysis of the data obtained from the inspections and possible modifications or actions on the asset stock.

In order to assess a segment's risk of failure, its probability of failure should be combined with its consequences of failure (Le Gauffre *et al.*, 2007; Tran *et al.*, 2009). The probability of failure could be assessed by deterioration models. However, determination of consequences of failure for a sewer segment is a complex procedure due to multidimensional impacts of a sewer segment failure.

According to Salman (2010), one of difficulties of assessing a segment's risk of failure is due to the high uncertainty involved in determining the direct costs associated with the emergency repair or replacement activities. Even higher uncertainty and difficulty exists in calculating the indirect costs of a sewer failure in monetary terms due to their intangible nature. The indirect costs of sewer failures include, but are not limited to, service disruptions, traffic delays (especially if the sewer is located under a roadway with high traffic volume), regulatory fines, environmental damage and public health problems. However, instead of determining consequences of failures in monetary terms, agencies may also develop consequences of failure indices in order to make comparisons and identify areas that will face a higher impact due to a potential failure. Consequently for simplicity reasons, this could be measured on a relative scale (Le Gauffre *et al*., 2007; Salman 2010).

McDonald & Zhao (2001) and Le Gauffre *et al*., (2004) introduce a set of criteria used to prioritize the inspection of sewer segments by taking into account the impact of segment failure on the surrounding environment. However, these approaches require the assessment of "impact factors" (vulnerability indicators, risk factors or consequences of failure) for which data is not always available (Halfawy *et al*., 2008; Hahn *et al*. 2002). Le Gauffre *et al*., (2004) describe how to construct an inspection criterion considering the impact of a segment's dysfunction on the environment. In the end, a multi-criteria analysis allows the stakeholders to define the appropriate considerations of their asset stock about these criteria by giving them appropriate importance weights.

Fuchs-Hanusch *et al*., (2012) by adopting a risk assessment methodology, a definition of relevant hazards causing "undesired events" responsible for deficits regarding functional requirements and by using a bivariate logistic regression analysis to derive the main influencing factors on the occurrence of these hazards aim at prioritizing the inspection needs. Moreover, they utilize a vulnerability analysis to quantify hydraulic driven consequences of undesired events. These methods were applied to one part of an Austrian sewer system and the results proofed that this risk-based approach improves significantly the inspection programs in terms of targeting the problematic segments.

Hintz *et al*. (2007) assessed criticality of segments on an ordinal grade of 1 to 3 based on pipe category (material), depth, proximity to environmentally sensitive areas and other critical assets, difficulty of access, system redundancy, likelihood of failure and flow rates/capacity.

Debères *et al*., (2011) and Ahmadi *et al*., (2011) used the risk-based approach in order to prioritize the rehabilitation needs within the master plan of the territory of Caen-la-Mer urban community, France by adopting the methodology developed in Le Gauffre *et al*., (2004). Their results show that small portion of total inspected length of sewers requires a rehabilitation action. For being more cost-effective, in the end, Ahmadi *et al*., (2011) propose to target segments with highest risk of failure within future inspection surveys.

3.2.2. Other existing approaches for inspection programs

Berardi *et al*., (2009) developed a sewer inspection prioritization program using the multi-objective genetic algorithm approach. The objective function consists of three terms: a) The cost of the inspection program 2) the expected cost of emergency repairs due to sewer blockages and 3) the expected cost of emergency repairs due to sewer collapses. The optimization process was repeated 5 times. Both the number of returned solutions (different inspection schemes) and selected segments vary considerably over the 5 runs (no consistent solution found). On the other hand, once an inspection program is chosen it is impossible to modify it due to an unexpected event, and in this case a whole new inspection plan should be regenerated. Furthermore, the approach is very time-consuming and requires a huge amount of data.

Baur and Herz (2002) develop an inspection methodology by forecasting condition of sewers from a small sample of inspected sewers of Dresden city, Germany by using a deterioration model as a means of predicting the most probable date of entering critical condition grade from sewer influential factors such as material, period of construction, location, network type, diameter and gradient.

3.2.3. Inspection techniques

Sewer inspections are carried out in order to identify structural and/or hydraulic problems (Salman 2010; Rahman and Vanier 2004). Ratliff (2003) classifies inspection process into three levels of assessment according to the inspection technique's capability and the required information. These three levels are:

- Field reconnaissance: which consist of collecting data of manholes and segments and the possibility of passing inspection equipment or staffs through manholes.
- Internal inspection: aiming at assessing the internal condition of segments. Some of available inspection techniques for no-man entry segments are:
    a. **Closed circuit television (CCTV)**: CCTV inspections have been employed for over 20 years (WERF 2007) and are one of the most widely used inspection methods (Najafi 2005). A CCTV equipment comprises a color CCTV camera and a lighting system installed on a wheeled carriage with a remote controlled function operated by the CCTV operator. Small modules are moved through the sewer by a winch and pulley system. However, bigger segments allow self-propelled modules to be used. The images captured by the CCTV camera are sent to the control center for image storage.

       CCTV inspections, however, provide only an assessment of the internal surface. It is recommended to flush and cleanse the segments prior to inspection to remove surface encrustations and bio film layers (WERF 2007; Najafi 2005).

Types of defects that can be identified by using CCTV inspections include cracks, sags, infiltration and inflow rates, tree roots, penetrating connections, presence of grease and offset joints. European standard EN13508-2 "condition of drain and sewer systems outside buildings, Part 2: Visual inspection coding system" provides an exhaustive list of possible identifiable defects by CCTV inspections.

However, this method is not applicable for pipes with high flow rates and allows only identification of defects located above the flow line. Furthermore, the accuracy of these inspections depends on the objectivity, experience, concentration of the operator and the quality of television pictures (Allouche and Freure 2002, Chae and Abraham 2001, Dirksen *et al*., 2013).

b. **Sonar or ultrasonic inspections**: according to WEF/ ASCE (2009) a sound wave is sent to the surface of the segment and the reflected wave is analyzed in order to identify defects. This inspection method could be combined with CCTV inspections to perform a comprehensive inspection. This method is not usable on segments made in brick. However, this method allows:
   i. Identify the existing defects
   ii. Determine the amount of debris
   iii. Determine the profile of segment
c. **Laser-based scanning method**: this method has been used to evaluate both the shapes of segments and the types of defects they contain (Hibino *et al*., 1994). This method is restricted to the part of the sewer above the waterline. An additional advantage of this method is that information from the laser scanners is readily recorded and analyzed by computer, substantially reducing operator errors. The equipment is more expensive that a CCTV system but the reduction in operator consumed-time could balance this difference. According to WEF/ASCE (2009), this method can be used to:
   i. Determine the shape and cross sectional area of segments
   ii. Determine the defects on the segment's wall
   iii. Determine the amount of debris
   iv. Determine the capacity before and after cleansing
   v. Determine the quality of the lining work.
d. **Sewer scanning and Evaluation Technology (SSET)**: is a flexible non-destructive evaluation data acquisition tool. The SSET removes the

deficiencies of the CCTV to provide the engineer with more and higher quality information. This is accomplished by utilizing scanner and gyroscope technology. The information produced provides the engineer the ability to see the total surface of the pipe from one end to the other. The scanned image is then digitized so that a color-coded computer image can be printed out. Defects are illustrated by a designated color code. A written description of each defect is produced at the appropriate location along the pipeline (Najafi 2005). According to WEF/ASCE (2009), the advantages of SSET over CCTV are as follows:

i. Interpretation error of the engineer (operator) is minimized by using the same light and angle settings
ii. Position of the camera and inclination of the pipe are recorded
iii. Image of the entire surface wall is obtained
iv. Camera should not be stopped necessarily in order to record defects
v. Data obtained from the inspection is analyzed by using computer software

e. **Smoke testing**: is used to identify faulty or illegal connections to gravity sewer and storm water systems. Another use of this test is to identify locations where infiltration and inflow occurs. Smoke from either smoke bombs or a liquid smoke system is forced into the system at manholes using special designed fans. It escapes from the system at house vent pipes, illegal connections and faulty ones, allowing them to be identified. This test is inexpensive and provides a fast method for locating illegal and faulty connections but it is not applicable under active precipitation, high groundwater or frozen ground connections. However, testing large segments can be difficult due to the capacity of equipment.

f. **Other methods**: there are, however, other methods which could be used as inspection techniques such as: Ground-penetrating radar (Daniels, 2005), inside jacking test (Thepot, 2004), mechanical impedence (Şabanoviç and Ohnishi, 2011) and etc. Makar (1999), WERF (2007), McClosket (2003), Najafi (2005) and WEF/ASCE (2009) compare all these methods and Salman (2010) provides a recapitulative table for this comparison.

- Inspection evaluation which consists of the assessment process of inspected segments according to a condition grading protocol.

3.2.4. Role of available data in inspection programs

The efficiency of sewer inspection programs (see chapter 4 for the definition of this notion) improves if the inspection survey targets segments in failure state. Hence, it seems that accurate predictions of the current and future condition of sewers are crucial for effective decision-making (Baik *et al*., 2006). These predictions can be obtained from deterioration models (Tran *et al.,* 2009).

Various types of deterioration models could be found within scientific literature (c.f. section 3.2.4.2.). However, despite the development of various deterioration models, attention is still focused on the type of deterioration model used and the influence of available data on the predictive power of these models has gone unstudied.

3.2.4.1. Factors affecting sewer deterioration

The structural deterioration of sewer segments is complex (WERF 2010). The failure mechanism is an event that causes the segment to reach one of combined strength and serviceability limit states (Farshad 2006). Strength limit state defines a condition at which the strength of the segment is reached. Examples of this state may be by loss of water tightness, and loss of stiffness (WERF 2010). The serviceability limit state defines a condition at which a particular function of the pipe is no longer fulfilled. Examples of this state may be larger deformations, buckling, clogging, abrasion and local damages (Farshad 2006).

According to Rahman and Vanier (2004), the process of failure of a sewer segment could be decomposed into a three-phase development of structural defects in conjunction with a concept of random damage events:

- The first step is initiated from minor defects such as cracks or leaking joints that are possibly caused by poor and improper construction or installation methods.

- The second step is the extension of the initial deterioration of step 1 in different rates depending on a combination of attacks such as external and internal load, chemical corrosion, erosion and ground loss.
- The final step usually occurs through probabilistic damage events such as nearby excavation or excessive load which is not easy to forecast. However, it is possible to judge if a segment has deteriorated sufficiently for failure or not.

According to the European standard EN 13508-2 and WRc's studies, the commonly found structural defects, as it is mentioned above, include cracks, fissures, fractures, deformation or shape distortion, hole, corrosion etc.

Makar and Kleiner (2000) and WERF (2009) state that the deterioration process leading to segment failure varies with segment's material but the rate of deterioration depends on exposure to different environments and operational conditions. In other words, each segment has a different rate of deterioration because contributing factors or explanatory factors vary among them (Tran *et al*., 2009).

WEF/ASCE (2009) and WERF (2009) describes the role of segment's material on the deterioration process. For example according to the former, concrete sewer segments are susceptible to corrosion due to the hydrogen sulfide present in the flow. Ablin and Kinshella (2004) describe the process of corrosion within sanitary systems.

Davies *et al*., (2001a) provide a review of previous researches on the factors that have an influence on the structural condition of sewer segments. They categorize these factors into three distinct groups (table 3-1). Considering this study along with Chughtai and Zayed (2008) and Müller (2002) almost every combination of explanatory variables (influential factors) can be found.

Table 3-1. Factors influencing on the deterioration rate of sewers (adapted from Davies *et al*., 2001a)

Construction factors	Local external factors	Other factors
Installation method	Surface use	Sewage characteristics
Standard of workmanship	Surface loading (including construction traffic)	Use of inappropriate maintenance methods
Sewer size	Surface type	Asset age
Sewer depth	Traffic characteristics	Sediment level
Bedding material and type	Water main bursts/leakage	surcharge
Sewer material	Ground movement	
Joint type and material	Maintenance of the other buried services	
Segment length	Groundwater level	
connections	Infiltration/exfiltration	
	Soil/backfill type	
	Root interference	

Davies *et al*., (2001a) conclude that it would be difficult to comprehensively support identified factors with data routinely collected by utilities. Furthermore, statistical variable selection procedures might result in different models if explanatory variables (influential factors) are strongly correlated. Nevertheless, there seems to be a core set of parameters that frequently show high relevance across various studies (Scheidegger *et al*., 2011). Müller (2002) identified these as age, size, depth and material. Dirksen and Clemens (2008) conclude that age is the most relevant explanatory factor for sewer decay. Ariaratnam *et al.,* (2001) analyze effects of diameter, material type, age, average depth of cover and waste type (debris). In contrary, Davies *et al*., (2001b) did not find a relevant influence of age on the deterioration of sewers. However, in general, age seems to be of very high importance as it contains two important types of information:

- Date of construction
- Indirectly the possible method and quality of construction.

3.2.4.2. Deterioration models

As it is mentioned above, current and future condition of assets are crucial for all aspects of a proactive asset management approach. Current assets' conditions are often assessed by inspection techniques cited above. However, this assessment has been carried out only for a small fraction of the assets. Consequently, the current conditions of un-assessed assets as well as future condition of all assets need to be predicted. Therefore, accurate predictions of the current and future condition of sewers are crucial for effective decision-making (Baik *et al*., 2006). In addition, in some cases it is also interesting and/or necessary to investigate the relationship between independent factors and dependent variable which is condition grade (Ana *et al*., 2009; Davies *et al*., 2001b). These predictions can be obtained from deterioration models (Tran *et al.,* 2009).

It is obvious that selecting the appropriate deterioration modeling technique will increase the accuracy of predictions. Various deterioration models have been developed and used in the literature, which utilize in general samples of evaluated CCTV inspection. Morcous *et al*., (2002) classifies the existing deterioration models into three categories: 1) deterministic models (linear and exponential regression) 2) probabilistic models (Markov chain, logistic and ordinal regression, linear discriminant analysis) and 3) artificial intelligence-based models (case-based reasoning, fuzzy set theory and neural networks). Salman and Salem (2012) provide a summary of models used to model sewer deterioration process. They also divide the deterioration models into two distinct main groups: 1) cohort-level and 2) individual section-level models.

WERF (2009) classifies the existing modeling approaches into:

- Deterministic models: in which relationships between influential factors and condition grade are assumed to be certain;
- Statistical models: in which the occurrence of failure over time is treated as a stochastic process and is represented by an appropriate probability distribution;
- Physical probabilistic models: which is based on an understanding of the physical processes that lead to asset failure while accounting for realistic uncertainty;
- Soft-computing or artificial intelligence models: in which model structure is determined by the data available and no prior relationship is assumed.

Table 3-2 summarizes the characteristics of these models and provides some examples of each approach.

Table 3-2. Deterioration models

Model	characteristics	Specific approach	Applications
Deterministic	relatively simple to develop and apply need historical data only applicable on homogeneous cohorts relying on a number of simplifying assumptions not accounting for the uncertainty associated with asset deterioration process	Empirical Physical	Kleiner and Rajani (2001) Randal-Smith *et al*., (1996) Rajani and Kleiner (2001)
Statistical	needing an amount of data in order to predict the behavior of deterioration process (model's parameters) applicable to any assets cohort level and/or individual section level model's parameters allowing to describe the influence of each influential factor	Failure event data-based Service lifetime-based Cohort-survival Ordinal regression Logistic regression Markov chain Bayesian networks	Kleiner and Rajani (2004); Rajani and Kleiner (2007); Salman and Salem (2012); Herz (2002); Baur and Herz (2002); Davies *et al*., (2001b); Koo and Ariaratnam (2006); Tran *et al*., (2009); Wirahdikusumah *et al*., (2001); Baik *et al*., (2006)
Physical probabilistic	Useful where no historical data is available Attempting to account for realistic uncertainty Normally use Monte Carlo simulations and some assumptions on the probability distributions of variables Cohort level and/or section level	Monte Carlo simulation Structural reliability theory	Davies *et al*., (2007); Rajani and Makar (2000); Mogolia *et al*., (2008);
Soft-computing	Data-driven model Needing an amount of data to be trained Black box effect Overall design of model depends on the user's knowledge (Number of nodes in different layers etc.)	Artificial Neural Networks (ANNs) Fuzzy logic	Tran *et al*., (2009); Achim *et al*., (2007); Rajani *et al*., (2006)

Salman (2010), Trans (2007) and WERF (2009) describe the advantages and inconvenient of using each of these methods. Salman and Salem (2012) by comparing various section-level deterioration models, conclude that logistic regression is the best method amongst them. In contrary, Trans (2007) showed that using neural networks is more efficient by dividing his small database, containing 417 segments, into two groups, first group for training or calibrating models and second one as test sample. However, because of his small database, the use of the logistic regression or all other types of statistical deterioration models is problematic and needs further investigations (for considering problematic issues regarding the calibration and interpretation of these deterioration models see section 3.4. and chapter 6).

Furthermore, Syachrani (2010) by developing a dynamic deterioration model shows that the patterns of deterioration among segments within a network vary depending on their physical and operational conditions. The utilization of location related attributes (e.g. land use) and other data (root problem, grease problem) are shown to be helpful to efficiently categorize segments into several clusters representing different patterns of deterioration. Furthermore, he compared three deterioration modeling techniques and showed Bayesian model has higher accuracy over regression and neural networks models.

Therefore, by considering all these studies, there is a competition between statistical models with soft-computing ones.

3.2.4.3. Historic records

A massive data collection effort is required to populate an asset register with accurate and meaningful data including physical information, age, condition, level of service, etc. (cf. 3.2.4.2; Syachrani 2010). According to Ana and Bauwens (2010), one of the main problems with the current deterioration models is that most of them are too rigid since they are developed for an ideal environment. Some models require extensive databases that are often not available at a utility level and some others use only small data sample which makes it difficult to generalize the results for the entire sewer network. We will readdress this issue in section 3.4 as well.

Another main problem regarding data available within utilities is a lack of historical records in terms of repair, renovation or replacement works that have been carried out on a network. Scheidegger *et al*., (2011) quote that calibrating a deterioration model on the basis of such data leads to an overestimation of segments' lifespans. However, datasets without historical records are common in practice because the main interest is often given to the current state of sewer segments.

Egger *et al*., (2013) quote that lack of the following information hinders the calibration of sewer deterioration models:

- When a segment is replaced, the corresponding records are discarded from the database and are replaced by the new information.
- Condition ratings of renovated or repaired segments before such actions are generally overwritten by a new and usually better condition.
- The renovation or repair of a segment is not recorded though it is assigned to a better condition grade.

They also suggest that calibration of a deterioration model without taking into account the historic of rehabilitation actions either lead to a "survival selection bias" resulting to longer segments' lifespans or "a simple calibration bias" as the condition of a segment is improved without recording the improvement. They address this issue by combining a probabilistic deterioration model with a simple rehabilitation model reasoning only on condition grade of segments and not on all random-base causes of rehabilitation such as roadworks. This combined model is then applied on 2 datasets: 1) on a well-defined synthetic generated dataset constructed by NetCoS (Scheidegger *et al*., 2011) and 2) on a real sewer network. Finally, the main outcome of their study is that, by taking into account all these considerations, the combined model compensates the bias induced by a lack of historical data.

However, they conclude that the estimation of their model parameters by frequentist inference (maximizing the likelihood of model) was not successful which brought them to reasoning on expert knowledge. On the other hand, another assumption of this reasoning is that the deterioration of sewer segments does not differ fundamentally between similar sewer networks in similar regions. This assumption is in contradiction with other researches proposing that expert opinions have shown important differences coming from similar utilities (Dirksen *et al*., 2013; Werey *et al*., 2008).

Moreover, in their model, apart from factor age, other factors (c.f. section 3.2.4.1) are not considered. Therefore, they do not take into account the replacement of a segment with a new diameter or material in their model. This latter is being practically done in France and in the U.S. considering the elimination of all segments in asbestos cement (EPA 2010). Finally, the main outcome of their study is that, by taking into account all these considerations, the combined model compensates the bias induced by a lack of historical data.

3.3. Assessment of a segment into a condition grade

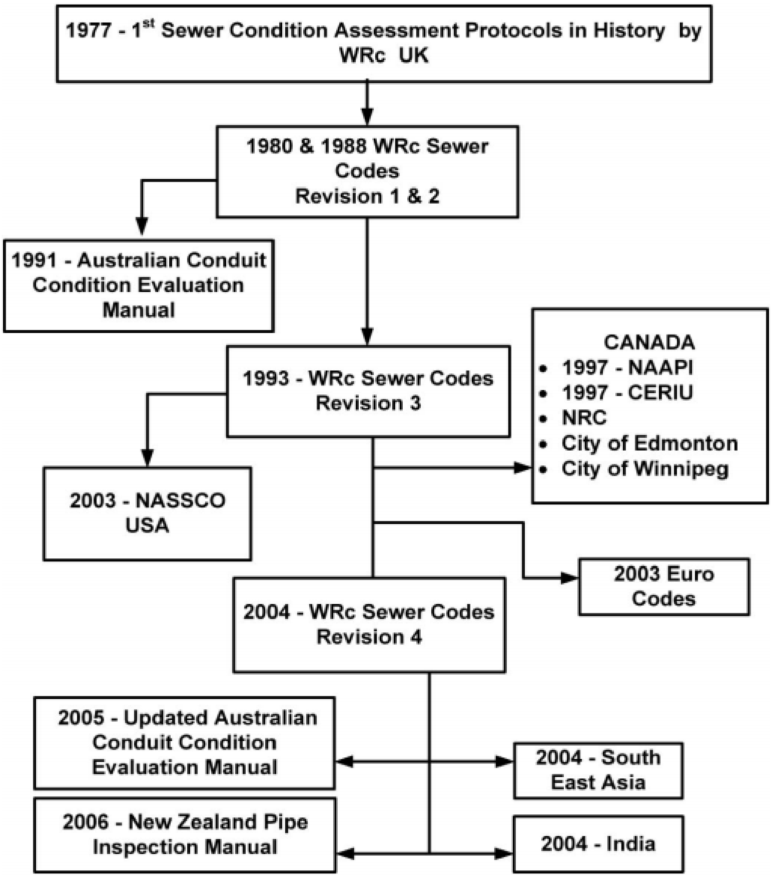
Sewer condition classification or condition rating has become of paramount importance for sewer asset management to ascertain critical information (Thornhill and Wildbore 2005). Condition rating is an evaluation of the infrastructure's current physical state versus its newly-constructed state supposed to be the perfect condition (Mehle *et al*., 2001). The main purpose of conducting condition ratings on assets is to determine their remaining useful life (WERF 2010).

Condition assessment looks at the current condition on an asset and establishes a benchmark for prioritizing maintenance and rehabilitation activities (Rahman and Vanier 2004). According to the latter, this task is a primary activity in implementing a successful sewer asset management program. However, in order to overcome huge maintenance backlogs and to improve the condition and performance of sewer systems, a unified condition assessment protocol is essential (Rahman and Vanier 2004). Currently, condition assessment of sewer networks is mainly based on CCTV or walk-through inspections (Wirahdikusmah *et al*., 1998).

3.3.1. Condition assessment protocols

According to Chughtai *et al.*, (2011), the historical background of the development of sewer condition assessment protocols goes back to 1977 when for the first time, sewer defect codes were developed by the Water Research Centre (WRc). Since then, several condition classification protocols have been developed throughout the world. Figure 3-4 shows an overview of this historic development adapted from Thornhill and Wildbore (2005).

Figure 3-4. Historical background of sewer condition assessment protocols (adapted from Thornhill and Wildbore 2005)



3.3.1.1. WRc

The WRc protocol divides sewer defects into two major categories: structural and operational. The evaluation of these defects (i.e. number and severity) leads to the assessment of the structural and operational condition of the pipeline. In addition to the structural and operational defects, WRc addresses some additional features such as construction defects.

These defects are used to identify the encountered and preexisting construction features for connections and joints, manholes and linings (Chughtai *et al.*, 2011).

To calculate a pipe's condition, sewer defects need to be ranked in some order of severity. Table 3-3 provides the most common defects considered by WRc protocol which could occur in a segment. Each defect can be classified as light, moderate or severe according to its size, number, shape and orientation. However, the classification of the defect and its severity can vary between inspectors (CCTV or walk-through operators) depending to their experience and the quality of images provided.

Table 3-3. Common defects according to WRc

Structural defects	Operational defects
Crack	Root
fracture	Debris
Deformation	Encrustation
Joint defects	Penetrating connection
Collapse	Infiltration
Break	
Sag	
Surface damage	
Corrosion	
Hole	

On the basis of the characteristics of defects, an overall sewer internal condition grade (CG) for the whole segment is identified by a number from 1–5 (WRc 2004) as illustrated in table 3-4. The CG for a segment is a numeric rating applied to visual images of sewer and is determined by a defect score calculation that is based on various defects in a pipe segment. The value of each defect (i.e. weight) determines the impact of the defect on the service life and performance of the sewer segment. The total score represents the summations of all deduct values in the segment, whereas the peak score represents the highest deduct value. The mean of the defect scores per meter of pipeline reflects its overall condition (NZWWA 2006).

Table 3-4. Severity condition grades for WRc protocol

Condition grade	Description	Peak structural score found in a segment	Peak operational score found in a segment
1	Acceptable condition	< 10	< 1
2	Minimal collapse risk but potential for further deterioration	10-39	1-1.9
3	Collapse unlikely but further deterioration likely	40- 79	2-4.9
4	Collapse likely in near future	80- 164	5-9.9
5	Collapse imminent or collapsed	165 and higher	> 10

Allouche and Freure (2002) carried out a survey on the maintenance and management practices for storm and sanitary sewers in order to determine the use of condition assessment techniques in Canadian municipalities serving 5.2 million people (17% of Canadian population). The results showed that about 70% of municipalities used the WRc protocol. However, most large Canadian municipalities in the survey were directly or indirectly using WRc assessment methods or had developed their own condition assessment method based on WRc guidelines.

3.3.1.2. NRC

National Research Centre has published *Guidelines for Condition Assessment and Rehabilitation of Large sewers* (Zhao *et al*., 2001). The condition assessment methods are limited to large sewers defined as sewers of 900 mm diameter and higher. These guidelines define various structural and operational defects along with a number of severity levels.

The deduct values assigned for these defects are then used to calculate the condition grade of the segment. The deduct values range from 1 to 20 for structural defects and from 1 to 10 for operational defects (table 3-5). Six condition grades are proposed for both structural and operational conditions. Table 3-5 provides the definition of these condition grades as well.

Table 3-5. Condition grades for NRC protocol

Condition grade	Description	Structural deduct value ranges	Operational deduct value ranges
5	Failed or failure imminent	20	9-10
4	Very poor condition High structural risk	15-19	7-8
3	Poor condition Moderate structural risk	10-14	5-6
2	Fair condition Minimal structural risk	5-9	3-4
1	Good condition	1-4	1-2
0	Excellent condition	0	0

3.3.1.3. NASSCO

National Association of Sewer Service Companies (NASSCO) developed PACP (Pipeline Assessment and Certification Program) protocol which is based on WRc protocol explained earlier (NASSCO 2004).

The PACP coding system categorizes defects and features into five sections: continuous defect coding, structural defect coding, operational and maintenance coding, construction features coding, and miscellaneous features coding. For each type of defect, the PACP uses a combination of capital letters to describe the type of defect and a number to rank the severity of the defect. Defect codes are recorded on a standardized form along with pertinent system data including defect type, continuous distance of the defect, severity, size, circumferential location (clock location), joint number, image/video reference number, and comments.

This protocol considers only internal segment conditions and has condition grades for both structural and operational defects. Segment ratings are based on the number of occurrences for each defect's condition grade and are calculated separately for both structural and operational defects for each segment. Each segment will be assigned into a grade based on the number of occurrences of each graded defect. The graded defect is multiplied by the number of occurrences, and this equals the segment grade. The overall pipe rating is calculated by adding all of the segment grades. The structural defects are added separately from the operational grades. Therefore, each pipeline receives two separate grades. Table 3-6 provides the definition of these condition grades and the estimated time to failure for each condition grade according to NASSCO (2004).

Table 3-6. Condition grades for PACP protocol

Condition grade	Description	Description	Estimated time to failure
5	Immediate attention	Defects requiring immediate attention	Has failed or will likely fail within the next 5 years
4	Poor	Severe defects that will become Grade 5 defects within foreseeable future	5 to 10 years
3	Fair	Moderate defects that will continue to deteriorate	10 to 20 years
2	Good	Defects that have not begun to deteriorate	20 years or more
1	Excellent	Minor defects	Unlikely in the foreseeable future

3.3.1.4. DWA-M 149-3, Germany

The German methodology DWA-M 149-3 (2011) is based on the European inspection standard EN 13508-2.

The methodology evaluates the operational and structural condition of a sewer pipe considering the most severe defects, the defect density and the influence of boundary conditions (e.g. groundwater level, sewer depth).

Single defects are evaluated according to their relevance in respect to three fundamental requirements: Leak tightness (L), Stability (S) and Operational safety (O) separately. They are rated according to 5 interim condition classes from 0 (very severe deficit with danger in delay) to 4 (minor deficit) for each requirement. For each sewer and each requirement, a condition score is calculated based on the most severe interim condition class and on the density of defects. Boundary conditions can be considered by adding an extra factor for each requirement.

Finally, the individual evaluations for each requirement are merged into a coefficient of rehabilitation need based on the most severe requirement. The coefficient of rehabilitation need can be transformed into a sewer condition class that defines the time-horizon for rehabilitation actions (for more details see Kley *et al*., 2013).

3.3.2. Uncertainty within assessment process

Although visual inspection is widely used in order to assess assets' conditions, many sources of uncertainty exist within the process of constructing rehabilitation programs from the results of investigations. Four major types of uncertainties related to visual inspection data can be remarked as follows (cf. chapter 5):

- Inspecting a segment and saving the existing defects by codes defined by sewer inspection codes such as PACP, EN 13508-2 or WRc by CCTV operator. Dirksen *et al*., (2013) illustrates how uncertainty is introduced into inspection reports by CCTV operators.
- The conversion process of codes into a score for each segment. This conversion depends on the severity, shape and length of defects which could be taken into account by a peak score, total score or mean score calculated for a given segment (WRc 2004; Chughtai *et al*., 2011).
- Assigning a condition grade to a segment from its score considering some thresholds. In other words, the main question here is that for two segments evaluated by WRc protocol, if the corresponding scores are successively 80 and 164, both of them are assigned into G4 (table 3-4). However, the nature and severity of defects could be very different.
- The fourth type of uncertainty is related to the table used to aggregate an indicator derived from visual inspection with another indicator such as risk factor or vulnerability of the environment (for assessing a rehabilitation criterion). Crossing indicator X (grade G2) with indicator Y (grade G3) to assess indicator Z may lead to two possible crisp aggregation results: G2 by an optimistic and G3 by a pessimistic view. Which option should be chosen? This type of issue is addressed in Le Gauffre and Cherqui (2009), which compares crisp and fuzzy operators.

Despite the development of various condition grading protocols, they all fail to address the following point:

- As it is said above, the condition assessment process contains some uncertainty sources. The condition assessment process for a segment could result in an over-estimation or an under-estimation of the real condition grade of the segment (see chapter 5). However, the above condition grading protocols do not allow the stakeholders to take into account their own specifications. Some stakeholders are very sensitive to the under-estimation of their assets' conditions (Werey *et al.*, 2008; Cherqui *et al.*, 2008). Therefore, the need of a more specific condition grading protocol is felt which could take into account the sensitivity of utility managers and stakeholders to this specific issue.

3.4. Need of assessing an asset stock from a sample

According to WERF (2009) a major step of strategic asset management is to establish the current overall condition of assets as a means of prioritizing and forecasting rehabilitation needs as well as economic analysis of different possible scenarios for future.

According to the American Society of Civil Engineers, the current average condition grade of the national wastewater asset stock is *D* below the average of *D+* for all infrastructure systems in the United States using a simple grading scale from *A* to *F* (ASCE 2013). These evaluations have been being carried out since 1988. According to the former, the overall condition of assets aims at evaluating the infrastructure's in-time or near future physical condition. Along with the last evaluation, according to the prediction carried out by the U.S. Environmental Protection Agency (EPA) by 2020, 44% of sewer segments will be in a condition grade that necessitates an imminent action (Allbee 2005). These evaluations are based on samples from the whole U.S. asset stock.

In fact, at the moment small number of utilities has completely inspected their asset stocks. Therefore, the use of a sample from an asset stock in order to calibrate deterioration models, to study scenarios about future and to draw appropriate conclusions seems mandatory. This sample should reflect, however, the characteristics of the asset stock in-question in the best manner. By definition, a sample which is an appropriate image of an asset stock is the representative sample of this asset stock (Cochran 1977; Lohr 2010).

Furthermore, Le Gauffre *et al.*, (2004); Le Gauffre *et al.*, (2007) propose an evaluation methodology in order to assign segments into an ordinal grade of G1-G4 (G1 the best condition and G4 the worst condition) based on CCTV inspections (c.f. Chapter 5). This methodology requires two parameters: 1) stakeholders' intentions about under-estimation or over-estimation of segments' conditions related to the uncertainty sources within CCTV evaluation process and 2) the overall condition grade of the asset stock for which having a representative sample of the asset stock or having some reliable guess about the overall condition of the asset stock is mandatory.

Baur and Herz (2002) forecast condition of sewers from a supposed representative sample of inspected sewers of Dresden city, Germany. Transition functions (based on cohort-survival functions) from one class into another condition grade were empirically derived from this sample. With these transition functions, the most probable date of entering critical condition grade can be forecast from sewer influential factors such as material, period of construction, location, network type, diameter and gradient. A procedure is proposed for scheduling the inspection dates for sewers which have not yet been inspected and for those which have been inspected before.

The use of a representative sample of an asset stock is not just limited to calibration of deterioration models or to provision of the overall condition of an asset stock. In addition, in chapter 4 of this thesis, by using the deviance statistic (a likelihood ratio test), we propose a method allowing the establishment of the list of most informative single factors in terms of inspection program efficiency (finding segments in failure state). This method can be applied on a representative sample of the asset stock (containing segments' influential factors such as age, material etc. and their condition grades) in order to give utilities important recommendations about data acquisition plans responding to following question: *"what data to gather considering its importance and cost of acquisition"*.

However, within the scientific literature dedicated to the asset management, authors developed and calibrated deterioration models without paying attention to the impact of used sample on the outcomes. The main interest of doing so was to test and compare the deterioration models (Trans 2007) or to develop a risk-based rehabilitation approach (Salman and Salem 2012).

Baur and Herz (2002) calibrated their cohort transition functions by using only 2.7% of the total length of Dresden network as a representative sample of this network. In addition, they used quota sampling method to carry out this representative sample. Quota sampling depends strongly on the user judgment and is not a probability sampling method (Cochran, 1977).

Hence, attention should be paid to following issues:

(1) The fact that we use just a sample of the asset stock. In other words, we should consider the following question: "Are our conclusions drawn from a sample generalizable to the whole asset stock?". For example Davis *et al.*, (2001b) by calibrating the binary logistic regression on a sample of the asset stock of Thames Water, UK, conclude that age, depth, traffic type and road type are not significantly different from zero. They also comment that age is marginally insignificant and they remove it from the analysis which in next level, could be applied on the whole asset stock.

(2) What if we have not observed all patterns existing within the asset stock in our sample? For example, assume that the proportion of segments in brick within an asset stock is small and within our available sample any segment does not represent segments in brick. Hence, we cannot draw any conclusion for segments in brick from this specific sample in first place before passing into the next level which is generalization of findings to the whole asset stock. For instance no study tackles with the problem of representativeness of the used samples in order to calibrate deterioration models.

(3) However, depending on the deterioration model used and available sample, the calibration process may be problematic. Concato *et al.* (1993) quote that multivariable methods of analysis have been suspected of producing problematic results if too few outcome events are available relative to the number of independent variables analyzed in the model. For example in the case of using the logistic regression, three types of errors have been identified (Peduzzi *et al*., 1996): over-fitting (Type I error) occurs when too many variables (factors), some of which may be “noise,” are selected for retention in the final model; under-fitting (Type II error) occurs when important variables are omitted from the final model; and paradoxical fitting (Type III error) is produced when a particular factor is given an incorrect direction of association which is the opposite of the true effect. Therefore, studies suggest a minimum number of events (failure state) per variable (EPV) about 10-20 EPV (Harrell *et al*., 1985; Peduzzi *et al*., 1995) in order to correctly calibrate the logistic regression coefficients. Bagley *et al*., (2001) compare the existing studies in this field and conclude that it is recommended that authors, reviewers, and editors pay greater attention to guidelines concerning the use and reporting of logistic regression models.

In the medical sciences, logistic regression modeling is commonly used because of its ability to model dichotomous outcomes. Proper use of this powerful and sophisticated modeling technique requires considerable care both in the specification of the form of the model and in the calculation and interpretation of the model's coefficients (Bagley *et al*., 2001).

A regression model serves two purposes: (1) it can predict the outcome variable for new values of the predictor variables, and (2) it can help answer questions about the area under study, because the coefficient of each predictor variable explicitly describes the relative contribution of that variable to the outcome variable, automatically controlling for the influences of the other predictor variables (Courvoisier *et al*., 2011). They report an important study on the issue of the number of events per variable (EPV) in logistic regression modeling. The article clearly shows that EPV>10 is no guarantee for unbiased estimation of regression coefficients which is in direct contradiction by Harrell *et al*., (1985); Peduzzi *et al*., (1995) proposing 10 EPV to correctly calibrate the logistic regression coefficients.

Hence, in total, these issues could be regrouped as follows:

- How to draw a representative sample of an asset stock which reflects in the best and appropriate manners the characteristics of this asset stock?
- How to provide a reliable estimation of a specific property of the asset stock from this sample such as the proportion of segments in failure state (condition grade 5 according to WRc for example)?
- Calibration of multivariate models seems somehow tricky. Therefore, what is the impact of used sample on the calibration outcomes of these multivariate models?

In chapter 6, first, we will introduce the probabilistic sampling methods for drawing a representative sample of an asset stock by considering different sampling methods: Simple random sampling, Proportional and Neyman (optimum) allocations in stratified sampling (literature review of this part is provided in chapter 6 for keeping the coherence of the text). Then, we will study the influence of sampling methods on the provision of a reliable estimation of a specific property of the asset stock. At the end, the impact of sampling methods (and consequently samples drawn) on the calibration process and on the interpretation of its results will be discussed.

3.5. Summary of literature review

The future-retrofitting benefits of asset management approaches have attracted so much attention by researchers, government agencies and public and private sector managers. Though asset management definition remains a controversial issue between researchers but all of them agree on the fact that having a well-defined asset management approach is a 3 dimensional challenge: economic, social and environmental.

A vital component of these approaches is the condition assessment of sewer segments. Inspection programs should be established in order to detect and evaluate deterioration of assets. Development of an effective inspection program is centered on knowing when, where and how to inspect. High numbers of deficient segments resulting from poor past maintenance activities and budget restrictions necessitate the development of prioritization tools to address inspection needs of the segments with the highest risk of failure.

The efficiency of sewer inspection programs improves if the inspection survey targets more problematic segments. Hence, it seems that accurate predictions of the current and future condition of sewers are crucial for effective decision-making. These predictions can be obtained from deterioration models. Though various types of deterioration models could be found within scientific literature, attention is still focused on the type of deterioration model used and the influence of available data on the predictive power of these models has gone unstudied.

Furthermore, once segments are inspected, they should be evaluated by an assessment protocol which remains identic for all segments. Despite the development of various condition grading protocols, they all fail to address the following point:

- The condition assessment process contains some uncertainty sources. The condition assessment process for a segment could result in an over-estimation or an under-estimation of the real condition grade of the segment, However, the existing condition grading protocols do not allow the stakeholders to take into account their own specifications about the under-estimation or over-estimation of their assets' conditions. Therefore, the need of a more specific condition grading protocol is felt which could take into account the sensitivity of utility managers and stakeholders to this specific issue.

In fact, at the moment small number of utilities has completely inspected and evaluated their asset stocks. Therefore, the use of a sample from an asset stock in order to calibrate decision-support models as deterioration models, to study scenarios about future and to draw appropriate conclusions for the whole asset stock seems mandatory. This sample should reflect, however, the characteristics of the asset stock in-question in the best manner.

However, by considering the scientific literature dedicated to asset management of water and wastewater industry, following problematic issues are identified:

- How to draw a representative sample of an asset stock which reflects in the best and appropriate manners the characteristics of this asset stock?
- How to provide a reliable estimation of a specific property of the asset stock from this sample?
- Calibration of multivariate models seems somehow tricky. Therefore, what is the impact of used sample on the calibration outcomes of these multivariate models?

Hence, the first objective of this dissertation is to propose a framework for inspection programs from which the influence of available data on the efficiency of these programs could be studied. Afterwards, we aim at testing a specific condition grading procedure which could take into account issues mentioned earlier. At the end, we will discuss some methods of drawing a representative sample of a given asset stock as well as the role of drawn sample in the calibration outcomes of a multivariate model.

In next chapter, the elaboration of inspection programs and the influence of available data on them will be studied. In chapter 5, we explain the proposed condition grading protocol and some sensitivity analyses of its parameters are provided. In chapter 6, we will study the sampling methods in order to draw a representative sample of an asset stock as well as the influence of available sample on the outcomes of a multivariate model. At the end, chapter 7 will provide the final conclusions and some new fields for further researches.

3.6. References

1. Ablin, R. L., and Kinshella, P. (2004). "Dude, where's my pipe – accelerated corrosion rate threatens Phoenix sewers." Conf. Proc., Pipeline Engineering and Construction: What's on the Horizon?, ASCE, Reston, VA., 1 – 8.
2. ABS (2004), American Bureau of Shipping. Guidance Notes on Reliability-centered maintenance, July 2004.
3. Achim, D., Ghotb, F., and McManus, K.J. (2007) Prediction of Water Pipe Asset Life Using Neural Networks, *Journal of Infrastructure System*, 13(1), 26-30.
4. Ahmadi, M., Cherqui, F., De Massiac, J.C., Le Gauffre P. (2013). "From sewer inspection programs to rehabilitation needs: research and results related to data quality and availability with the support of numerical experiment". *European Journal of Civil Engineering* (In press).
5. Ahmadi M., Debères P., De Massiac J.C., Le Gauffre P., Cherqui F., Werey C. (2011). "Using INDIGAU asset management tool for prioritizing rehabilitation programs of sewage network of Caen-la-Mer". The International Conference on Water and Wastewater, with focus on Privatization and Benchmarking, Tehran, Iran.
6. Allbee, S., (2005). "America's pathway to sustainable water and wastewater systems". Water Asset Management International, pp. 9-14.
7. Allouche, E.N. and P. Freure (2002). "Management and Maintenance Practices of Storm and Sanitary Sewers in Canadian Municipalities", ICLR Research, Paper Series No. 18.
8. Ana Jr., E., and Bauwens, W. (2007). "Sewer Network Asset Management Decision-Support Tools: A Review." International Symposium on New Directions in Urban Water Management, UNESCO, Paris.
9. Ana, E., and Bauwens, W., (2010). "Modeling the structural deterioration of urban drainage pipes: the state-of-the-art in statistical methods". *Urban Water Journal*, 7(1), 47 – 59.

10. Ana, E., and Bauwens, W., (2010). “Modeling the structural deterioration of urban drainage pipes: the state-of-the-art in statistical methods”. *Urban Water Journal*, 7(1), 47 – 59.
11. Ana, E., Bauwens, W., Pessemier, M., Thoeye, S. Smoders, S., Boonen I. and De Gueldre G. (2009): “An investigation of factors influencing sewer structural degradation”, *Urban Water Journal*, 6:4, 303-312.
12. Ariaratnam, S. T., El-Assaly, A. and Yang, Y., (2001). “Assessment of infrastructure inspection needs using logistic models”. *Journal of Infrastructure System*. 74, 160–165.
13. ASCE (2013). “Report card for America’s infrastructure.” <http://www.infrastructurereportcard.org/> (Jun. 20. 2013)
14. ASCE/WEF (1994), “Existing Sewer Evaluation and Rehabilitation,” ASCE Manual No.62 or WEF MOP FD-6, Edition 2, 1994.
15. Bagley SC, White H, Golumb BA, (2001). “Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain”. *Journal of Clinical Epidemiologic*; 54:979-85.
16. Baik, H. S., Jeong, H. S. D. and Abraham, D. M., (2006). “Estimating transition probabilities in Markov chain-based deterioration models for management of wastewater systems”. *Journal of Water Resource Planning & Management*, ASCE, 132(1), 15–24.
17. Baur, R. and Herz, R. (2002). “Selective inspection planning with ageing forecast for sewer types”. *Journal of Water Science and Technology*. 46, 4-7, 389-396.
18. Berardi, L., Giustolisi, O., Savic, D.A. and Kapelan, Z., (2009). “An effective multi-objective approach to prioritisation of sewer pipe inspection”. *Water Science and Technology,* 60 (4), pp. 841–850
19. Burns, P., Hope, D., Roorda, J. (1999). “Managing Infrastructure for the Next Generation,” *Automation in Construction*, Elsevier, 8(6), 689-703.
20. CEN, 2003, EN 13508-2: “Condition of drain and sewer systems outside buildings, Part 2: Visual inspection coding system”, European Norms.
21. Chae, M., Abraham, D. (2001). “Neuro-Fuzzy Approaches for Sanitary Sewer Pipeline Condition Assessment.” *Journal of Computing in Civil Engineering*, 15(1), 4-14.
22. Cherqui, F., Werey, C., Ibrahim, M., Le Gauffre, P., (2008). “CCTV inspection of sewer segments: calibration of performance indicators based on experts’ opinions”. Proceedings of the 11th International Conference on Urban Drainage, Edinburgh, UK, 31 Aug.-5 Sept., 10 p.
23. Chughtai, F. and Zayed, T. (2011). ”Integrating WRc and CERIU Condition Assessment Models and Classification Protocols for Sewer Pipelines”. *Journal of Infrastructure System*. 17(3), 129–136.
24. Chughtai, F., and Zayed, T. (2008). “Infrastructure Condition Prediction Models for Sustainable Sewer Pipeline.” *Journal of performance of constructed facilities*, ASCE, 22(5), 333-341.
25. Cochran, William G. (1997). “Sampling Techniques, 3rd Edition”. John Wiley, New York.
26. Concato, J., Feinstein, AR., Holford, TR., (1993). “The risk of determining risk with multivariable models”. *Journal of Annals of Internal Medicine*; 118: 201-210.

27. Daniels, D. (2005). "Ground penetrating radar". Encyclopedia of RF and Microwave Engineering.
28. Davies, J. P., Clarke, B. A., Whiter, J. T. and Cunningham, R. J., (2001a). "Factors influencing the structural deterioration and collapse of rigid sewer pipes". *Urban Water Journal*, 3(1–2), 73–89.
29. Davies, J. P., Clarke, B. A., Whiter, J. T., Cunningham, R. J. and Leidi, A., (2001b). "The structural condition of rigid sewer pipes: A statistical investigation". *Urban Water Journal*, 3(4), 277–286.
30. Davis, P., Burn, S., Moglia, M., and Gould, S. (2007). "A Physical Probabilistic Model to Predict Failure Rates in Buried PVC Pipelines", *Reliability Engineering and System Safety*, 92, 1258-1266.
31. Debères P., Ahmadi M., De Massiac J.C., Le Gauffre P., Cherqui F., Werey C., Nirsimloo K. (2011). « Deploying a sewer asset management strategy using the indigau decision support system". 4th international leading-edge conference on Strategic Asset management, LESAM, Mulheim, Germany.
32. Dirksen, J. and Clemens, F.H.L.R., (2008). "Probabilistic modeling of sewer deterioration using inspection data". *Water Science and Technology*, 57 (10), 1635–1641.
33. Dirksen, J., Clemens, F. H. L. R., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K., Snaterse, C. T. M., (2013). "The consistency of visual sewer inspection data", *Structure and Infrastructure Engineering*. Vol. 9, Is. 3, pp. 214-228. (DOI: 10.1080/ 15732479.2010.541265).
34. DWA-M 149-3, (2011). "Advisory Leaflet DWA-M 149-3 - Conditions and Assessment of Drain and Sewer Systems Outside Buildings – Part 3: Condition Classification and Assessment", ISBN: 978-3-942964-11-1. German original published 2007, English translation published 2011.
35. Egger, C., Scheidegger, A., Reichert, P., Maurer, M., (2013). "Sewer deterioration modeling with condition data lacking historical records", *Water Research*. (DOI: 10.1016/j.watres.2013.09.010).
36. EPA (2002): U.S. EPA SSO Fact Sheet – "Asset Management for Sewer Collection Systems", Office of Wastewater Management, 833-F-02-001, April 2002.
37. EPA (2010): "State of Technology for Rehabilitation of Wastewater Collection Systems", United States, Environmental Protection Agency, Office of Research and Development, EPA/600/R-10/078, July 2010.
38. Farshad, M., (2006). "Plastic Pipe Systems: Failure Investigation and Diagnosis," ISBN 1-85617- 496-4, Elsevier UK.
39. Fuchs-Hanusch, D., Friedl, F. and Kogseder, B. (2011). "Effect of Seasonal Climatic Variance on Water Main Failures in Moderate Climate Regions." Conference Proceedings of LESAM 11, Leading Edge in Strategic Asset Management 27.-29. Sep. 2011, Germany
40. Fuchs-Hanusch, D., Friedl, F., Möderl, M., Sprung, W., Plihal, H., Kretschmer, F., Ertl, T. (2012). "Risk and Performance Oriented Sewer Inspection Prioritization". World Environmental and Water Resources Congress. ASCE.

41. Gemora, I., (2003). PACP: "Pipeline Assessment and Certification Program". Proc. Int. Conf. on Pipeline Engineering and Construction, ASCE, Reston, VA, 822–825.
42. Hahn, M., Palmer, R., Merrill, M., Andrew, B., Lukas, A., (2002). "Expert system for prioritizing the inspection of sewers: Knowledge base formulation and evaluation". ASCE, Journal of Water Resources Planning and Management, 128(2), 121-129.
43. Halfawy, M. R., Dridi, L. and Baker, S., (2008). "Integrated decision support system for optimal renewal planning of sewer networks". *Journal of Computing in Civil Engineering*, 22(6), 360–372.
44. Harrell, F.E., Lee, K.L., Matchar, D.B., Reichert, T.A., (1985). "Regression models for prognostic prediction: advantages, problems, and suggested solutions. Cancer Treatment Reports"; 69:1071–7.
45. Herz, R. (2002). "Software for Strategic Rehabilitation and Investment Planning". International Conference Computer Aided Rehabilitation of Water Networks CARE-W. Dresden. 65-84.
46. Hibino, Y., Nomura, T., Ohta, S. and Yoshida, N, (1994). "Laser scanner for tunnel inspections, International Water Power and Dam Construction", June, IPC Electrical-Electronic Press, London, England.
47. Hintz, A. M., Barnes, D., and Millar, D.C. (2007). "Establishing a collection system baseline condition assessment program one step at a time." Proc., Pipelines 2007: Advances and Experiences with Trenchless Pipeline Projects, ASCE, Reston, VA.
48. IAM/BSI, (2004): (PAS-55). "Asset management, Part 1: Specification for the optimized management of infrastructure assets. Institute of Asset Management & British Standard Institution" – Business Information, 24 p.
49. IPWEA, (2006), "International Infrastructure Management Manual", Institute of Public Works Engineering Australia, Sydney.
50. Kleiner, Y. and Rajani, B.B. (2004). "Quantifying effectiveness of cathodic protection in water mains", *Journal of Infrastructure Systems*, 10(2), 43-51.
51. Kleiner, Y., and Rajani, B. (2001). "Comprehensive review of structural deterioration of water mains: statistical models." *Urban Water*, 3(3), 131-150.
52. Kley, G., Kropp, I., Schmidt, T., Caradot, N. (2013). "Review of available technologies and methodologies for sewer condition evaluation". Berlin, Germany. http://www.kompetenz-wasser.de/fileadmin/user_upload/pdf/forschung/SEMA/D11_SEMA_Review_of_technologies_and_methodologies_for_sewer_condition_evaluation.pdf
53. Koo, D.-H. and Ariaratnam, S.T. (2006) Innovative Method for Assessment of Underground Sewer Pipe Condition, Automation in Construction, 15, 479-488.
54. Le Gauffre, P., Cherqui, F., (2009). Sewer rehabilitation criteria evaluated by fusion of fuzzy indicators. LESAM 2009 – 3rd IWA Leading Edge Conference on Strategic Asset Management. Miami, U.S.
55. Lohr S. L. (2010). "Sampling: Design and Analysis". Cengage Learning. Boston, USA.
56. Makar, J. and Kleiner, Y. (2000). "Maintaining water pipeline integrity", AWWA Infrastructure Conference and Exibition, Baltimore, Maryland, USA, March 12-15, 2000.

57. Makar, J.M., (1999). “Diagnostic techniques for sewer systems”. *Journal of Infrastructure Systems,* 5 (2), 69–78.
58. Marlow, D. and Burn, S., (2008). “Effective use of Condition Assessment within Asset Management”, *JAWWA*, 100(1), 54-63.
59. Marlow, D., Davis, P., Trans, D., Beale, D., Burn, S., Urquhart, A., (2009). “Remaining asset life: a state of the art review”, WERF – Water Environment Research Foundation, IWA Publishing, London, UK.
60. Marlow, D., Heart, S., Burn, S., Urquhart, A., Gould, S., Anderson, M., Cook, S., Ambrose, M., Madin, B. and Fitzgerald A., (2007). “Condition assessment strategies and protocols for water and wastewater utility assets”, WERF - Water Environment Research Foundation, IWA Publishing, 476 p.
61. McCloskey, A. (2003). “A three level strategic approach to investigating and evaluating wastewater networks.” Conf. Proc., New Pipeline Technologies, Security, and Safety, ASCE, Reston, VA., 1580 – 1591.
62. McDonald, S. and Zhao, J., (2001). “Condition assessment and rehabilitation of large sewers”. International conference on Underground Infrastructure Research, University of Waterloo, Waterloo, Ontario, June 10-13, 2001, 361-369.
63. Mehle, J.J., O’Keefe, S.M. & Wrase, P.E. (2001). “An Examination of Methods for Condition Rating of Sewer Pipelines”. Center for Development of Technical Leadership, University of Minnesota.
64. Merrill, S., Lukas, A., Palmer, R.N., Hahn, M.A., (2004). “Development of a tool to prioritize sewer inspections”. Technical report, WERF - Water Environment Research Foundation, 123 p.
65. Moglia, M., Davis, P., and Burn, S. (2008). “Strong Exploration of a Cast Iron Pipe Failure Model”. *Reliability Engineering & System Safety*, 93, 885-896.
66. Morcous, G., Rivard, H. & Hanna, M. A., (2002), "Case-Based Reasoning System for Modeling Infrastructure Deterioration", *Journal of Computing in Civil Engineering*, ASCE, Vol. 16, No. 2, pp. 104-114.
67. Müller, K., (2002). “Development of a Generally Applicable Methodology for Selective Primary Inspection of Sewer Pipes and House Connecting Pipes. Part A: Scientific Investigations”. (Entwicklung eines allgemein anwendbaren Verfahrens zur selektiven Erstinspektion von Abwasserkanalen und Anschlussleitungen. Teil A: wissenschaftliche Untersuchungen. Abschlussbericht). Final report. Chair and Institute of Environmental Engineering, RWTH Aachen, Aachen, Germany (in German).
68. Najafi, M. (2005). “Trenchless Technology-Pipeline and Utility Design, Construction, and Renewal,” Water Environment Federation (WEF Press), Alexandria, VA.
69. NASSCO, (2001). “Pipeline assessment and certification program manual”, Pipeline Assessment and Certification Program (PACP), Second Edition Reference Manual, NASSCO.
70. NZWWA: New Zealand Water and Waste Water Association (NZWWA). (2006). New Zealand pipe inspection manual, 3rd Ed., NZWWA, Wellington, New Zealand.
71. Peduzzi, P., Concato, J., Feinstein, A.R., Holford, T.R., (1995). “The importance of events per independent variable (EPV) in proportional hazards regression analysis: II.

Accuracy and precision of regression estimates". *Journal of Clinical Epidemiologic*; 48: 1503-1510.

72. Peduzzi, P., Concato, J., Kemper, E., Holford, TR., Feinstein, A.R., (1996). "A simulation study of the number of events per variable in logistic regression analysis". *Journal of Clinical Epidemiologic*; 49: 1373–1379.
73. Rahman, S. and Vanier, D.J., (2004). "MIIP Client Report: An Evaluation of Condition Assessment Protocols for Sewer Management", National Research Council of Canada, Ottawa.
74. Rajani, B. and Makar, J. (2000). "A Methodology to Estimate Remaining Service Life of Grey Cast Iron Water Mains", *Can J Civil Eng.*, 27, 1259-1272.
75. Rajani, B., Kleiner, Y., (2001). "Comprehensive review of structural deterioration of water mains: physically based models". *Urban Water* 3 (3), 151-164.
76. Rajani, B., Kleiner, Y., and Sadiq, R. (2006). "Translation of Pipe Inspection Results Into Condition Ratings Using The Fuzzy Synthetic Evaluation Technique", *Journal of Water Supply Research and Technology*-Aqua, 55, 11-24.
77. Rajani, B.B. and Kleiner, Y. (2007). "Quantifying effectiveness of cathodic protection in water mains: case studies", *Journal of Infrastructure Systems*, 13(1), 1-11.
78. Randall-Smith, M., Russell, A., and Oliphant, R. (1992). "Guidance manual for the structural condition assessment of trunk mains", WRc, UK.
79. Ratliff, A. (2004). "An overview of current and developing technologies for pipe condition assessment". New Pipeline Technologies, Security, and Safety: pp. 103-114. DOI: 10.1061/40690(2003)16
80. Şabanoviç, A., Ohnishi, Kouhei. (2011). "Motion control systems", Singapore: John Wiley & Sons (Asia) Pte Ltd.
81. Salman, B., (2010). "Infrastructure management and deterioration risk assessment of wastewater collection systems". Thesis (PhD), Univ. of Cincinnati, OH.
82. Salman, B., and Salem, O., (2012). "Modeling failure of wastewater collection lines using various section-level regression models", *Journal of Infrastructure Systems*, ASCE, 18(2), 146-154.
83. Scheidegger, A., Hug, T., Rieckermann, J. and Maurer, M., (2011). "Network condition simulator for benchmarking sewer deterioration models". *Water Research*, 45(16), 4983-4994.
84. Schulting, F.L. and Alegre, H., (2007). "Global developments of Strategic Asset Management". LESAM 2007 – 2nd IWA Leading Edge Conference on Strategic Asset Management, International Water Association. Proceeding on CD-Rom. Lisbon 17-19 oct.
85. Sinha, S., Angkasuwansiri, T., THomasson, R. (2010). "Phase 1: Development of standard data structure to support wastewater pipe condition and performance prediction", WERF – Water Environment Research Foundation, IWA Publishing, London, UK.
86. Syachrani, S., (2010). "Advanced sewer asset management using dynamic deterioration models", PhD dissertation, Oklahoma State University, U.S.

87. Thepot, O. (2004). “Prise en compte des caractéristiques en petites deformations des sols dans l’étude du comportement des collecteurs enterrés”. Thesis dissertation, Ecole des Ponts, Paris, France.
88. Thornhill, R., and Wildbore, P. (2005). “Sewer defect codes: Origin and destination.” (http://www.nassco.org/training_edu/pdfs/040105_utech.pdf ⟩.
89. Tran, D.H., Ng, A.W.M., and Perera, B.J.C. (2007). “Neural Network Deterioration Models for Serviceability Condition of Buried Stormwater Pipes”, *Engineering Applications of Artificial Intelligence*, 20, 1144-1151.
90. Tran, D.H., Perera, B.J.C. and Ng, A.W.M., (2009). “Comparison of structural deterioration models for stormwater drainage pipes”. *Comput. Aided Civ. Infrastruct. Eng*. 24 (2), 145–156.
91. Tran, H.D., (2007). “Investigation of deterioration models for stormwater pipe systems”. PhD dissertation. Victoria University.
92. Vanier, D.J., (2001). “Asset Management: A to Z”, NRCC/CPWA/IPWEA Seminar Series, Innovations in Urban Infrastructure, Philadelphia, PA, 1-16.
93. Water Research Center (WRc). (2004). “Sewerage rehabilitation manual”, 4th Ed., WRc, Wiltshire, U.K.
94. Werey, C., Cherqui, F., Ibrahim, M., Le Gauffre, P., (2008). « Sewer asset management tool: dealing with experts’ opinions”. In Cabrera & Pardo (ed.) Performance Assessment of Urban Infrastructure Services. Drinking water, waste water and solid waste. London, UK: IWA Publishing, 125-134.
95. Wirahadikusumah, R., Abraham, D.M., Iseley, T., and Prasanth, R.K., (1998), “Assessment Technologies for Sewer System Rehabilitation,”. *Automation in Construction*, Vol. 7, No. 4, pp. 259-270.
96. Wirahadikusumah, R., Abraham, M. D., and Iseley, T. (2001). “Challenging Issues in Modeling Deterioration of Combined Sewers”, *Journal of Infrastructure Systems*, ASCE, 7(2), 77-84.
97. Zhao, J., McDonald, S., Kleiner, Y. (2001). “Guidelines for Condition Assessment and Rehabilitation of Large Sewers”. Institute for Research in Construction, National Research Council Canada. Ottawa, Canada.

Chapter IV: Inspection programming and influence of data on its efficiency

4.1. Overview

This chapter is extracted from two papers successively entitled:

- *Influence of available data on sewer inspection program efficiency (Ahmadi et al., 2013a);*
- *Benefits of using basic, imprecise or uncertain data for elaborating sewer inspection programs (Ahmadi et al., 2013b).*

A vital component of proactive approaches in sewer asset management is the condition assessment of sewer segments which mainly involves visual inspection (Ana and Bauwens, 2010). According to Le Gauffre *et al*., (2007) a sewer segment corresponds to a length of several meters which is homogeneous in some characteristics (material, diameter, etc.). In general, it is also delimited by two successive manholes. Currently, Closed Circuit Television inspections (CCTV) are widely used to assess the condition of sewer segments. CCTV inspections provide an inventory of observed defects reported using a specific coding system such as the European standard EN 13508-2 (CEN 2003) or PACP in North America (Gemora 2003). The observed codes are then quantified and a condition state is assigned to the segment according to the condition grading protocol used (c.f. chapter 5).

In practice, CCTV inspections should target the sewers most likely to be in failure state (Davies *et al*., 2001a). Those sewers determined in failure state are then prioritized for immediate action.

The efficiency of sewer inspection programs improves if the inspection survey targets segments in failure state. Therefore, accurate predictions of the current and future condition of sewers are crucial for effective decision-making (Baik *et al*., 2006). These predictions can be obtained from deterioration models (Tran *et al*., 2009).

However, despite the development of various deterioration models, attention is still focused on the type of deterioration model used and the influence of available data on the predictive power of these models in terms of imprecision and incompleteness has gone unstudied.

4.2. Scope and objective of this chapter

One purpose of this chapter is to propose a systematic approach using a deterioration model (segment targeting model: figure 4-1) to predict the structural condition of sewers in order to improve the efficiency of sewer inspection programs (cf. 4.3.4). Another purpose of this chapter is to study the influence of the data available within a utility on the inspection programs. To this end, a complete database should be available. This will be degraded in order to introduce imprecision, incompleteness and uncertainty into the data available within a utility database (figure 4-1).

In the following section, our proposed systematic approach for inspection programs is described (simulation framework), some problematic issues regarding the influence of available data within a utility are addressed and finally some indicators for measuring the efficiency of inspection programs are introduced. These indicators make it possible to gauge the influence of one or more influential factors on the efficiency of inspection programs.

In section 4.4, a semi-virtual asset stock (SVAS) inspired from Salman (2010) and Salman and Salem (2012) is generated (process 1a, 1b and 1c in figure 4-1). This SVAS is then degraded to construct the utility database (UDB) in order to provide answers to the problematic issues introduced in section 4.3.2. This degradation allows us to study the effects of incompleteness (In), imprecision (Im) and uncertainty (Un) in existing data in UDB on the inspection programs. Finally the simulation results and a comparison of the efficiency of inspection programs will be discussed in section 4.5.

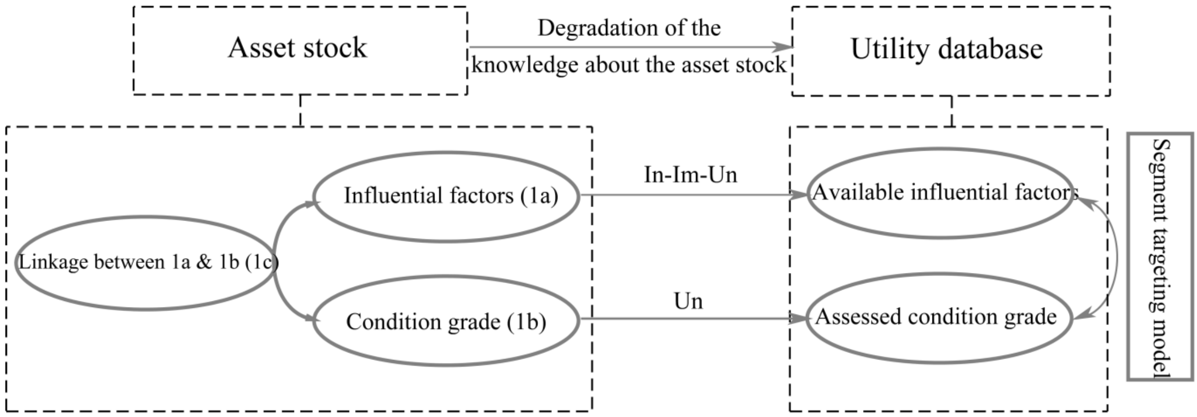


Figure 4-1. Difference between Asset stock and utility database. In: Incompleteness, Im: Imprecision and Un: Uncertainty.

4.3. Assessing the influence of data available within a utility database

This section describes the adapted framework for numerical simulations. Then, problematic issues are introduced regarding the impact of the incompleteness/imprecision of the utility database on inspection programs. Certain efficiency indicators are then developed in order to measure and to compare the efficiency of each inspection program.

4.3.1. Simulation framework

In this section, the framework of numerical simulations and its invariant parameters are presented. Assume that *p%* of the total length of an asset stock is in failure state. According to our knowledge of this asset stock, inspection programs can be divided into 4 main groups (figure 4-2):

(1) Perfect inspection program: inspecting p% of asset stock's length to find 100% of segments in failure state (theoretical and unachievable in practice);

(2) Inspection program with complete knowledge of sewers: all significant influential factors are available in database and their relationship with the segments' condition

grade is also known (unfeasible as models rely on the data obtained by inspection surveys);

(3) Annually progressive inspection program: this case shows the reality of inspection programs within utilities (scope of this chapter);

(4) Inspection program based on a random principle: no assumption on how the inspection should be carried out leads to a random program (inefficient).

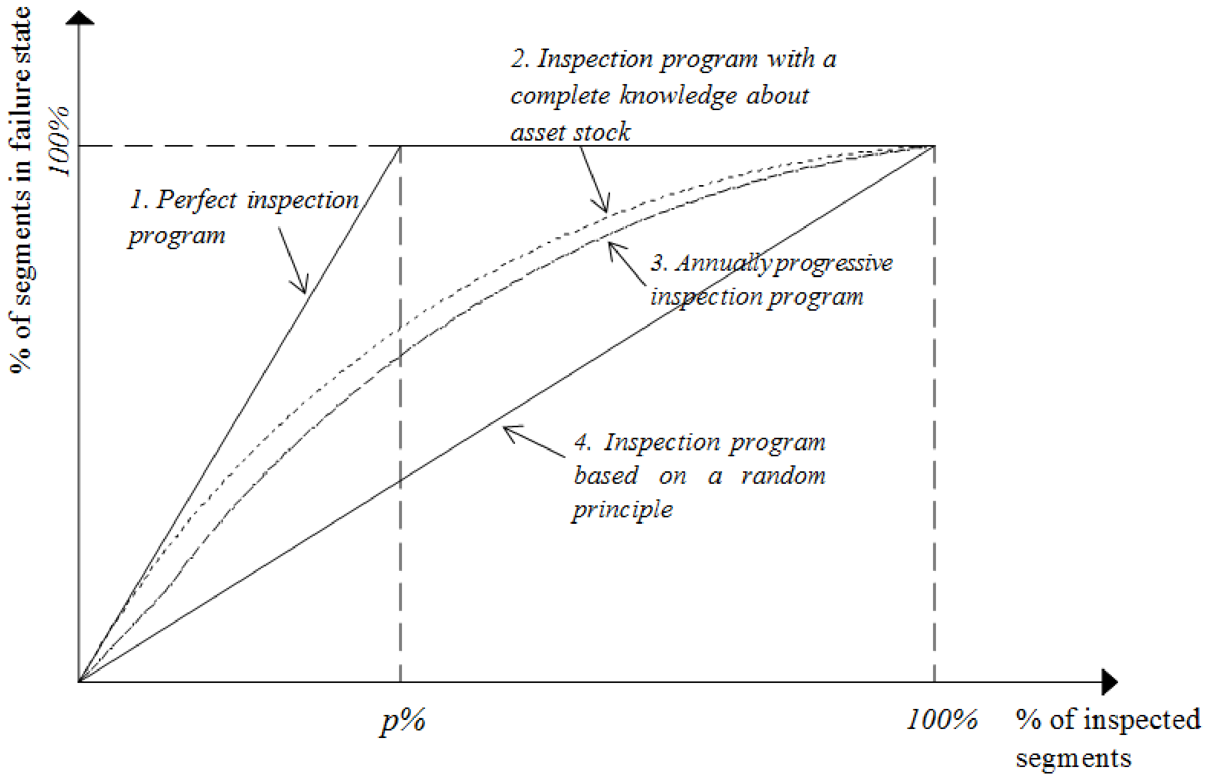


Figure 4-2. 4 main inspection program groups

Due to a lack of funds, each utility inspects a limited portion of its asset stock per year (AIR: annual inspection rate). Consequently, we suppose that it has an initial stock of inspected segments before implementing a new inspection program. If the influential factors for the segments are also available in the utility database, a probabilistic deterioration model can assess the segments' probability of being in failure state. Those segments with the highest probabilities are then selected to be inspected with respect to the AIR. We call this model

STM (segment targeting model). We use binary logistic regression as the key element of the STM (c.f. section 4.4.2.2).

We assume that the AIR is divided into two parts: 1) R: the portion of segments whose inspections are random, meaning that these segments are not chosen according to sewer asset management priorities (such as inspections motivated by roadworks, odor problems etc.) and 2) TA=AIR-R: the portion of segments which has the highest probability of being in failure state calculated by the STM (targeted segments). Once a segment is selected for inspection, it will be excluded from the list of uninspected segments. In other words, this approach offers only one selection possibility for a given segment.

Assume that this program is planned for *T* years. In the beginning (*t=0)*, the STM is calibrated using the initial stock of CCTV reports (observed condition grades of inspected segments) and available influential factors for these segments. Once the STM is calibrated, it will be used to calculate the probability of being in failure state for uninspected segments. By sorting them according to their probability of being in failure state, those with the highest probabilities are then selected to be inspected (until reaching TA% of the total length of the network). Then, the remaining R% is added randomly *(t=1)*. Except for the recalibration of the STM, the procedure remains unchanged for the coming years (until *t=T*). The recalibration process is carried out at the end of each year by considering the segments in the initial stock plus segments selected to be inspected during the previous years. The whole procedure is then repeated 100 times. Figure 4-3 demonstrates these steps while table 4-1 recapitulates the adapted invariant parameters of the framework.

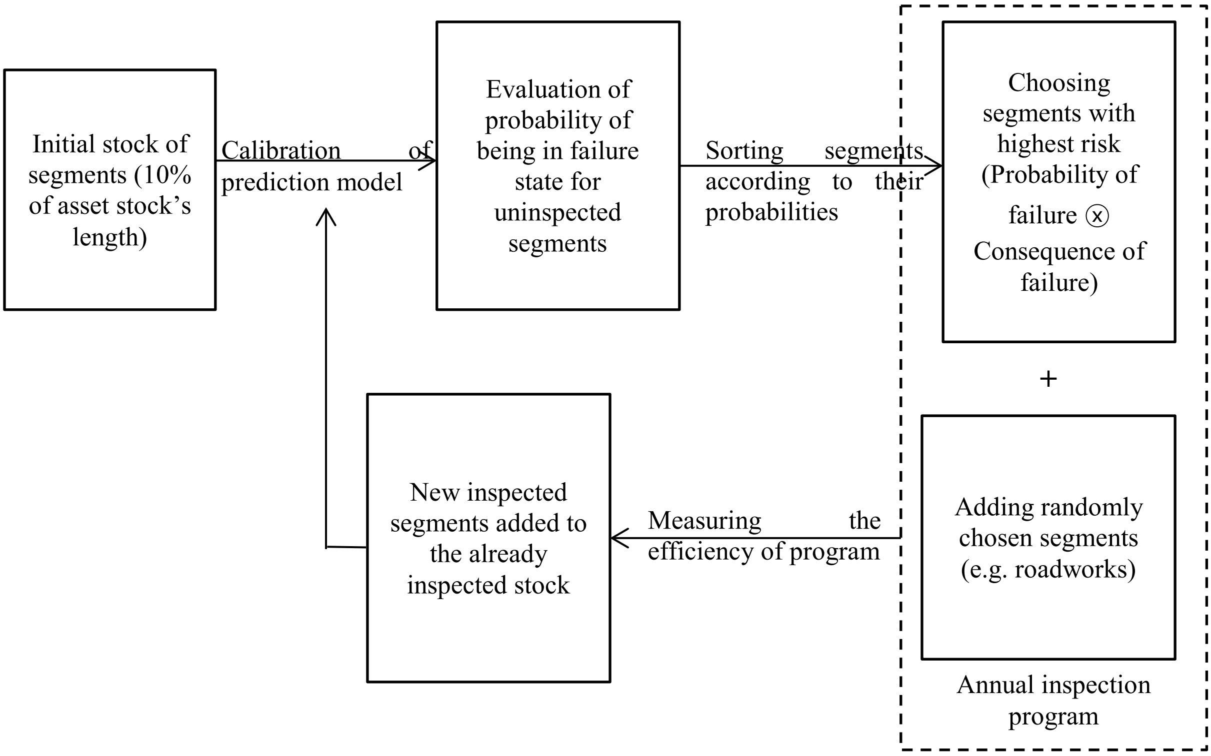


Figure 4-3: Framework of numerical simulations

Table 4-1: invariant parameters of the framework

Parameter	Assigned value
Annual inspection rate (AIR)	5% of total length of the network
Randomly-chosen segments rate(R)	3% of total length of the network
Targeted segments rate (TA)	2% of total length of the network
Duration of programs (T)	4 years
Quantity of initial stock of inspected segments	10% of total length of asset stock chosen randomly

4.3.2. Problematic issues: developing inspection programs

In this section, the influence of data available within a utility database on the efficiency of inspection programs in terms of a) data incompleteness b) data imprecision and c) data uncertainty will be questioned.

In other words, we will answer the following questions:

(1) Which factor is most informative? And how to establish the list of most informative single factors?

(2) Is it preferable to have imprecision instead of incompleteness within the utility database?

(3) How can the data most probably available within a utility be used to define an effective inspection program? For example, are data collected for hydraulic models of network such as segment's diameter, depth, gradient, sewer type and length, *hereafter* called basic data, sufficient and relevant enough for inspection programs?

(4) Can we use an auxiliary variable in order to compensate effects of missing data?

(5) Is it worth to accept a degree of uncertainty within data instead of not having them?

Table 4-2 summarizes these questions into four main categories. Numerical simulations are defined in order to respond to each category of questions. Each simulation consists of 100 Monte Carlo tries which are carried out by applying the framework described in section 4.3.1. The results are reported in section 4.5.

4.3.4. Indicators for measuring the efficiency of inspection programs

The efficiency of an inspection program is improved if segments in failure state are targeted by the inspection surveys. Breysse *et al.*, (2007) compares some simple inspection, maintenance and rehabilitation strategies by defining two indexes: 1) Technical performance index and 2) Technical and economic performance index. These indexes aim to analyze the profitability of investment in terms of both technical and economic performance. The cost of inspecting a segment can be considered either as a function of both segment size and segment length (Zhao 1998) or simply as a function of segment length (Yang and Su 2007). By adopting the latter definition, the overall efficiency of the inspection program considering both targeted and randomly-selected segments in year T_0 ($T_0 \in \{1,2,\dots,T\}$) is as follows:

$$\rho_{overall,t\leq T_0} = \frac{Total\ inspection\ cost\ of\ segments\ in\ failure\ state(t\leq T_0)}{Total\ cost\ of\ inspection\ program\ (t\leq T_0)} \quad (1)$$

Table 4-2: Inspection program problematic issues

Questions	Studying the influence of:	N.O. of simulations and their description
1- Which factor is most informative? How to establish the list of most informative factors?	data incompleteness	8 steps (36 simulations): By applying a forward selection process, each time, the most informative influential factor is added to the nested model until reaching the model with all factors (full-model). The informativeness of each remaining factor at each step is calculated using η (explained in section 4.3.4)
2- Is it better to have imprecision instead of incompleteness?	data imprecision and incompleteness	6 simulations: specifically focused on the imprecision of "age": All factors + age in form of a scale variable All factors + age in form of 2 age groups (≤50 or >50) All factors + age in form of 2 age groups (≤80 or >80) All factors + age in form of 4 age groups All factors + age in form of 7 urbanization periods (districts or neighborhoods) Without age (data incompleteness)
3- Are the *basic data* sufficient and relevant enough for inspection programs? Can we use an auxiliary variable in order to compensate the effect of missing data?	data incompleteness	25 simulations: by applying a forward selection process, each time, the most informative influential factor is added to the nested model until reaching the model with all factors (full-model). The informativeness of each remaining factor at each step is calculated by using η (explained in section 4.3.4).
4- Is it worth to accept a degree of uncertainty within data instead of not having them?	data uncertainty and incompleteness	10 simulations: specifically focused on the uncertainty within "age" variable (c.f. 4.4). The databases are defined as scenarios with different characteristics in terms of availability of each variable plus uncertainty within the factor "age" (c.f. 4.4).

$$\rho_{overall,t\leq T_0} = \frac{\sum_{t=1}^{T_0}\sum_{k\in I\_F_t} C.L_k}{\sum_{t=1}^{T_0}\sum_{j\in I_t} C.L_j} = \frac{\sum_{t=1}^{T_0}\sum_{k\in I\_F_t} L_k}{\sum_{t=1}^{T_0}\sum_{j\in I_t} L_j} \quad (2)$$

Similarly for targeted segments:

$$\rho_{tar,t\leq T_0} = \frac{\sum_{t=1}^{T_0}\sum_{k\in Tar\_F_t} L_k}{\sum_{t=1}^{T_0}\sum_{j\in Tar_t} L_j} \quad (3)$$

Where I_t is the set of all selected segments to be inspected in year t divided into Tar_t (set of targeted segments) and Ran_t (set of randomly-selected segments). I_F_t is the set of all inspected segments found to be in failure state in year t divided into Tar_F_t (set of targeted segments in failure state) and Ran_F_t (set of randomly-selected segments in failure state) and C is the cost of inspection per meter of length.

A normalized indicator is also defined which measures the ratio of segments found to be in failure state by the STM to the proportion of segments in failure state considering the whole asset stock (p):

$$\eta_t = \frac{\rho_{tar,t \le T_0}}{p} \tag{4}$$

Likewise, another indicator is introduced to compare two or more inspection programs independent of time *t:*

$\mu_{q\%}$= *percentage of segments found to be in failure state by inspecting q% of all segments* (5)

The values of T_0 and q are fixed respectively at 4 years and 30%. It should be noted that $\mu_{30\%}$ is the percentage of segments found to be in failure state at the end of first four years of program.

4.4. Data for numerical experiments: semi-virtual asset stock

The application of virtual case studies is commonly used in urban drainage to test measures, approaches or models (Urich *et al.,* 2010). Scheidegger *et al*., (2011) and Ana and Bauwens (2010) state that one of the major reasons sewer deterioration models fail to adequately predict the condition of sewers is the lack of complete and reliable datasets. Without reliable data, rehabilitation planning is little more than a guessing game. To this end, Scheidegger *et al.,* (2011) proposed a new network condition simulator that produces a synthetic population of sewer segments with a given condition-class distribution in order to test deterioration models. Urich *et al*., (2010) developed an agent based approach for generating virtual sewer systems to test hydraulic models and to optimize sewer placement. Nevertheless, their focus remains on the condition grade, localization and topology of segments and they do not tackle the problem of creating influential factors in virtual databases. Möderl *et al*., (2011) also generated 2,280 virtual water supply systems using the graph-theory-based Modular Design

System. The layout and the properties of these systems are representative of typical examples encountered in the real world. A comparison of the virtual sets with three real-world case studies shows similar characteristics between them.

For the purposes of this chapter, a semi-virtual asset stock (SVAS) has been created to study the influence of available UDB data on inspection programs. This SVAS is inspired by the asset stock of the metropolitan sewer district of Greater Cincinnati, USA (Salman 2010).

4.4.1. Sewer factors influencing segments' physical condition

Davies *et al.,* (2001a) provide a review of previous researches on factors influencing the structural deterioration of sewer segments and classify 25 influential factors into three main categories: 1) construction features (sewer size, depth, material & etc.) 2) local external factors (surface use, groundwater level & etc.) and 3) other factors (sewage characteristics, inappropriate maintenance methods). They also conclude that it would be difficult to comprehensively support identified factors with data routinely collected by utilities. Ariaratnam *et al.,* (2001) analyze effects of diameter, material type, age, average depth of cover and waste type, by applying a logistic regression model to predict the probability of a sewer segment being in failure state.

However, the inclusion of a large number of factors in models generates the following constraints (Tran *et al*., 2009):

- More time and effort required to build the models;
- Increase in computational complexity due to high-dimensional error surface;
- Increase in the sample size required to calibrate models (i.e. increase in the cost and resources required to obtain data).

4.4.2. SVAS generation methodology

A semi-virtual asset stock inspired by the asset stock of the metropolitan sewer district of Greater Cincinnati, USA (Salman 2010) is generated. The methodology of regeneration consists of:

(1) Regeneration of influential factors (process 1a in figure 4-1): age, material, size, depth, slope, road class, length and sewer type

(2) Assignment of condition grades to sewer segments from influential factors by using the logistic regression model provided by Salman and Salem (2012) (process 1b & 1c in figure 4-1).

The term "semi-virtual" reflects the fact that virtual sewer segments have been generated by using the influential factors distributions and logistic regression coefficients provided by Salman and Salem (2012) for the Cincinnati asset inventory.

4.4.2.1. Physical and environmental characteristics: regeneration of influential factors

Similar histograms to those provided by Salman (2010) are reproduced for all influential factors (table 4-3). However, minor changes are needed to convert from imperial units as used by Salman (2010) to metric units. These changes modified the following factors: size, depth and length.

The SVAS created contains 32,000 segments. The total length of the sewers in the SVAS is 737 km. The factor "age" includes the effects of construction period and ageing. The average age of the asset stock is 78.5 years. Approximately 60% and 24% of segments are made of vitrified clay and concrete respectively. About 40% and 20% of segments have a diameter equal to 200 and 300 millimeters respectively. The majority of segments are buried within a depth of 2 meters. The gradient for the majority of segments is less than 5%. Furthermore, about 70% of segments are buried under a street or an alley.

Table 4-3: Description of influential factors

Influential factor	Description	Nature of variable
Age	Difference between the installation year of a segment and the year of inspection	Scale
Material	1: Brick, 2: vitrified clay, 3: clay tile, 4: reinforced concrete or 5: concrete	Categorical
Size	Diameter of the segment in millimeter	Scale
Depth	Average depth of buried pipe from the ground level	Scale
Gradient	Vertical displacement of the segment per horizontal displacement in percentage	Scale
Length	Length of segment in meter	Scale
Sewer type	1: Combined or 2: sanitary	Categorical
Road class	0, if the segment is not located under any type of roadway 1, if the segment is located under a street or an alley 2, if the segment is located under an interstate highway, U.S and state route	Categorical

4.4.2.2. Condition grade of sewer segments

Ariaratnam *et al*., (2001), Davies *et al*., (2001b) and Salman and Salem (2012) used binary logistic regression to model the deterioration of sewer networks. Koo and Ariartnam (2006) also used this type of regression to assess the condition grade of large PVC lined concrete segments in the City of Phoenix, U.S.A. In this chapter, the coefficients of binary logistic regression provided by Salman and Salem (2012) are used to assign a condition grade to each segment based on its influential factors. The binary condition grade is expressed as structurally failed (CG=1) and not failed (CG=0).

Salman and Salem (2012) provide all coefficients derived from the asset stock of Greater Cincinnati in imperial units. To generate the SVAS, coefficients were adjusted in order to use the International System of Units. The condition grade of segments is assigned by considering their probability of being in failure state using these coefficients. However, we changed the intercept term in order to decrease the number of segments in failure state within the SVAS (p=6.61% defined in section 4.3.4, instead of 35.85% in the asset stock of Greater Cincinnati). This does not have any impact on the effects of influential factors on the segments' condition grade, as their coefficients remain unchanged. To this end, assume that *d* influential factors exist within the database. The general form of binary logistic regression is as follows:

$$L(x) = logit(Y=1) = \ln\left(\frac{P(Y=1|X_1,X_2,X_3,\ldots,X_d)}{1-P(Y=1|X_1,X_2,X_3,\ldots,X_d)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_d X_d \quad (6)$$

Where β_0= intercept term, $\beta_1, \beta_2, \ldots, \beta_d$= variable coefficients, $X_1, X_2, \ldots, X_d$= variables (influential factors), Y = binary dependent variable (condition grade).

The probability of being in failure state is then obtained by:

$$P(Y=1|X_1,X_2,X_3,\ldots,X_p) = \frac{1}{1+e^{-L(x)}} \quad (7)$$

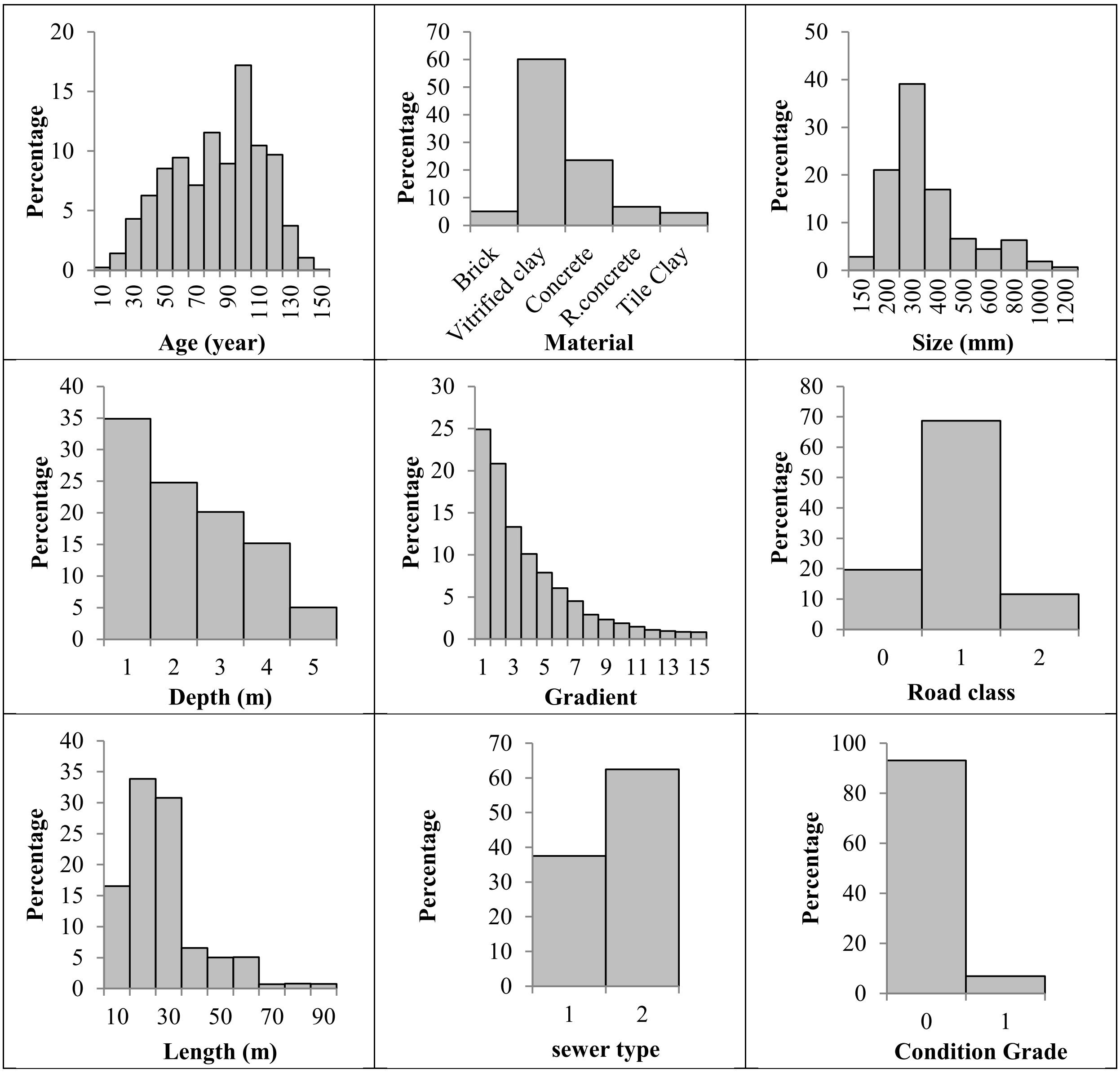


Figure 4-4: Histograms of influential factors and condition grade

Figure 4-4 shows the distributions of the influential factors and segment condition grades created. In addition, table 4-4 provides the binary logistic regression coefficients found after the SVAS was created (similar to those provided by Salman and Salem (2010)) using the maximum likelihood estimation method (Agresti 2002). Vitrified clay, sanitary and road class=2 values are fixed as reference categories for material, sewer type and road class variables. Table 4-4 also shows the Wald's significance test and *p*-values for each variable. All of the regression coefficients for scale variables are significantly different from zero at a significance level of 0.05 except for the depth coefficient. In addition, road class=1 is not significantly different from the chosen reference (road class=2) and they could be regrouped together. However, we kept all the variables in the model (depth is statistically significant from 0 at a 0.2 significance level).

Table 4-4: Binary logistic regression results

Factor	**Coefficient of created SVAS**	***p*-Value**
Intercept	-4.578	0.000
Age (year)	0.021	0.000
Material 1= Brick	-2.355	0.000
Material 3= Concrete	-0.627	0.000
Material 4= Reinforced Concrete	-1.463	0.000
Material 5= Clay tile	-1.204	0.000
Size (mm)	-0.0008	0.000
Depth (m)	-0.025	0.172
Gradient (%)	0.068	0.000
Road class=0	0.245	0.000
Road class=1	-0.03	0.779
Length (m)	0.0114	0.000
Sewer type= Combined	0.2754	0.000

4.4.3. Assigning sewer segments to urban districts

Arthur and Burkhard (2010) constructed a GIS based sewer condition model using a database of customer contacts and CCTV data. A major challenge in their study was estimating the age of each segment. It was inferred using the estimated age of surface developments by overlaying the GIS network with historical maps. They concluded that inferred asset age could be used to reliably highlight the assets most likely to fail.

In this paragraph, we introduce the notion of districts as an auxiliary variable derived from following hypotheses on existing factors in the SVAS:

(1) A district is an area of a city developed during a period of time characterized by a dominant activity (industrial, residential and etc.)

(2) The segments laid in a given district have almost the same age. In other words, the sewer network had been constructed either during surface development if the district is not very old or during a given period of time if the district is situated in the historical part of the city.

(3) The sewer type remains almost the same within the whole district; however this is not always true as an industrial-residential district may have two different types of network.

(4) A given material is dominant within each district. In other words, segments of a district are almost all made of a specific material.

We used the K-Nearest Neighbor (KNN) method (Fix and Hodges 1951) for introducing the notion of district into the SVAS as an auxiliary variable. The K-Nearest-Neighbor classification was developed from the need to perform discriminant analysis when reliable parametric estimates of probability densities are unknown. We can modify or bind the importance of each factor by applying a distance weighted KNN approach (Bailey and Jain, 1978) to the 3 factors cited in the hypotheses (age, material and sewer type).

Each segment is a point in 3D space. We decided to define 7 districts. By assigning seven points as seven centroids (each centroid represents a district labeled from 1 to 7), the distance between each point with each centroid is calculated. The centroid label nearest to each segment is then chosen as this segment's district. The only problem was the categorical nature of variables. This was solved by adapting a specific definition of distance between two points.

Consider $X_i = (x_{i,ST}, x_{i,A}, x_{i,M})$ as segment i ($i \in \{1,2,\dots,32000\}$) and $X_j = (x_{j,ST}, x_{j,A}, x_{j,M})$ as centroid j ($j \in \{1,2,\dots,7\}$). The distance between them is defined as follows:

$$d(X_i, X_j) = \sqrt{Im_{ST} \times Dis(ST) + Im_A \times Dis(A) + Im_M \times Dis(M)} \qquad (8)$$

Where:

Im_{ST}, Im_A and Im_M are respectively the importance weightings of sewer type, age and material (table 4-5). $Dis(ST), Dis(A)$ and $Dis(M)$ are the terms introduced in table 4-5 to measure the differences between these two points (i & j) for one factor at a time. They are selected in a manner to ensure somehow the homogeneity within most important factors: Age and Sewer type. The coordinates of all centroids are provided in table 4-6.

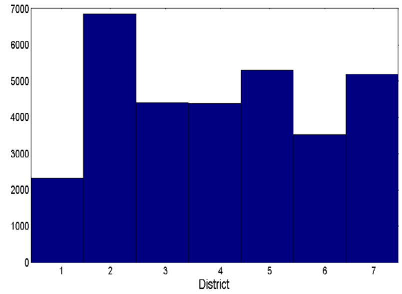
Table 4-5: The defined distance terms and importance weightings for each factor

Factor	**Nature**	**Term in defined distance**	**Importance weighting (*Im*)**	**Limit**
Sewer type	Categorical $x_{ST} = 1 \text{ or } 2$	If $x_{i,ST} = x_{j,ST}$ then Dis(ST)=0 Else Dis(ST)=1	36	-
Age	Scale but converted to categorical by considering each 10 years as one category labeled from 1 to 15 (between 1-10 as 1, 11-20 as 2, etc.)	Dis(A)= $(x_{i,A} - x_{j,A})^2$	1	If Dis(A)>36 then Dis(A)=36
Material	Categorical $x_M = 1 \text{ or } 2 \dots \text{or } 5$	If $x_{i,M} = x_{j,M}$ then Dis(M)=0 Else Dis(M)=1	9	-

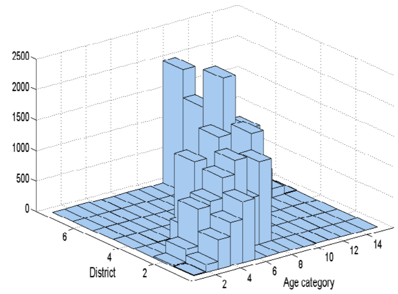
Table 4-6: chosen district centroids

District Centroid	**Age**	**Material**	**Sewer type**
1	30	2	1
2	50	2	2
3	70	3	2
4	80	2	1
5	100	2	1
6	120	3	2
7	130	2	2

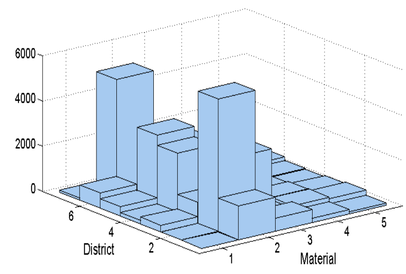
Figure 4-5 shows the histogram of districts and the factor distributions within each district.



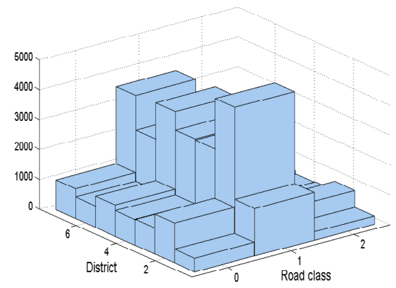
Histogram of segments within districts



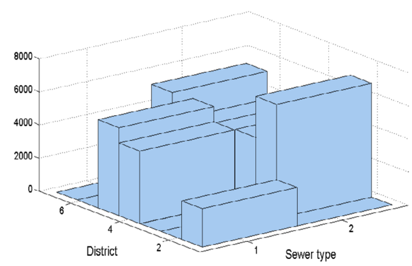
District vs. age



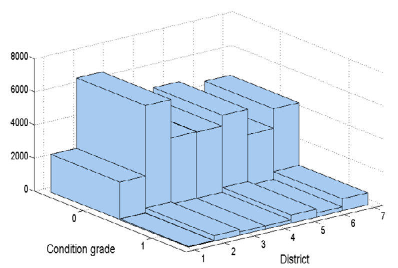
District vs. material



District vs. road class



District vs. sewer type



District vs. condition grade

Figure 4-5: Histogram of districts and their characteristics

4.4.4. Preliminary data analysis

A statistical analysis was conducted on the SVAS applying a correlation test for scale variables, cross table analysis for categorical variables and a backward selection method in order to find the importance rank of variables. These tests were performed using Matlab®.

4.4.4.1. Correlation test

This test is used to study the linear relationships between scale variables (Tran *et al.,* 2009). Pearson’s correlation coefficient measures the linear association level between two variables (Montgomery *et al.,* 2004). The values for this correlation test are between 0 and 1 ([0.85-1] strong linear relationship, [0.5-0.85] moderate linear relationship and [0-0.5] weak linear relationship). Pearson's correlation coefficient results follow Student's *t*-distribution. Therefore, the *t*-test is used to verify if the linear relationships are statistically significant or not. The results of this test are provided in table 4-7.

Table 4-7: Pearson's correlation test results

	Age	Size	Depth	Gradient	Length
Age	1.0000	0.0713*	-0.0124*	0.0047	-0.0006
Size	-	1.0000	-0.0085	0.0063	0.3340*
Depth	-	-	1.0000	0.0096*	-0.0001
Gradient	-	-	-	1.0000	0.0002
Length	-	-	-	-	1.0000

* Correlation is statistically significant at the 0.05 level

The results show that weak linear relationships exist between "age" and "size", "age" and "depth", "size" and "length", and "depth" and "gradient". The positive correlation between "size" and "length" is expected because the larger segments are often the main sewers which must be longer. Likewise, the positive correlation between "age" and "size" is also expected because during urban development the larger segments are laid first followed by smaller ones. However, the weak positive correlation between gradient and size, as well as the weak negative correlation between size and depth are not expected.

4.4.4.2. Cross-table analysis

The cross-table analysis is used to test the significance of categorical variables. This test measures the level of association between two categorical variables using the hypothesis of equality of cell counts across the table of two factors (Montgomery *et al.*, 2004; Tran *et al.*, 2009; Salman 2010).

Cramer (1994) introduces a Chi-square test statistic for a cross-table. The results of this test are shown in table 4-8. According to the results, all categorical variables have a statistically significant effect on the structural condition grade.

Table 4-8: Results of cross-table analysis

	Chi-square statistic	Degree of freedom	Significance
Material	334.4	4	0.00
Sewer type	22.5	1	0.00
Road class	21.2	2	0.00

4.4.4.3. Backward selection method

A backward selection method was carried out using the deviance statistic (McCullagh and Nelder 1989) and adapting a nested subset process (Guyon and Elisseeff 2003). The deviance statistic is a likelihood ratio test statistic (Selvin 1994).

Consider two binary logistic regression models: M_0 and M_1. If M_0 is the full model (with all available influential factors), the deviance of M_1 (reduced model or a nested model of M_0) from M_0 is calculated as follows:

$$D = -2(\log(p(M_1) - \log(p(M_0))) \quad (9)$$

This is −2 times the difference between the log-likelihood of the reduced model (M_1) and the full model (M_0) which reflects the difference in the fit of the two models. This follows an approximate chi-squared distribution with k-degrees of freedom (Model M_0 has k additional parameters compared to model M_1). The significance of the full model is evaluated based on the deviance statistic of the full model and the intercept-only model (Salman 2010) ($D_{full-model\,\&\,intercept-only\ model} = 1278.6 \rightarrow p - \text{value} = 0.000$) (table 4-9). The intercept-only model is the model without any influential factors. By excluding the variable which minimizes the deviance statistic at each step (having smallest effect on the model), the results of the backward selection method are shown in table 4-9.

The results show that all variables except depth are statistically significant as previously proven. According to table 4-9, the single most informative variable is age followed by material, gradient, length, sewer type, size, road class and finally depth.

We used the term “the single most informative variable” because there is a possibility that for example two variables with a weak explanatory power can possess a strong explanatory power when they are combined (Kleiner *et al.,* 2013). Therefore, possible interactions between variables should be studied in order to find the set(s) of variables which has (have) a strong explanatory power on the condition grade of segments. For example, the analysis may show that variable *depth* has a significant impact with some particular *materials* and/or with small *diameters*.

In this work, we constructed a semi-virtual asset stock from the information provided by Salman (2010) and Salman and Salem (2012) (cf. section 4.4). As they did not provide any information about these interactions, we were not able to consider them during the construction phase of the semi-virtual asset stock. Therefore, in this case, as these interactions are not introduced into the semi-virtual asset stock, their assessments may not bring any significant information. However, in real cases, it is mandatory to assess the most important pair, triplet … of variables according to recommendations given by Kleiner *et al.,* (2013).

Table 4-9: Backward selection method results and the results of significance tests for influential factors in binary logistic regression

Model consisting of:	**Exclusion of**	**-2log(likelihood)**	**Change in deviance**	**Degree of freedom**	***p*-value**	**Change in deviance/ degree of freedom**
Full-model	----	14850.6	---	---	---	---
Age, Material, Gradient, Sewer type, Length, Road class, Size	Depth	14852.47	1.87	1	0.172	1.87
Age, Material, Gradient, Sewer type, Length, Road class	Size	14870.73	18.26	1	0.0000	18.26
Age, Material, Gradient, Sewer type, Length	Road class	14891.48	20.75	2	0.0000	10.38
Age, Material, Gradient, Sewer type	Length	14918.9	27.42	1	0.0000	27.42
Age, Material, Gradient	Sewer type	14950.28	31.38	1	0.0000	31.38
Age, Material	Gradient	15054.03	103.75	1	0.0000	103.75
Age	Material	15538.45	484.42	4	0.0000	121.10
Intercept-only	Age	16129.22	590.77	1	0.0000	590.77

4.5. Results, discussion and recommendations

The results of the simulations defined in order to answer the questions asked in section 4.3.2 are reported in this section. These results are the average of 100 Monte Carlo tries for each simulation.

4.5.1. Influence of data incompleteness: how to establish the list of the single most informative factors?

4.5.1.1. Assessment of the inspection programs' efficiency

This section aims to study the influence of data incompleteness in the UDB. In order to respond to the questions asked in table 4-2, 8 modeling steps should be defined (36 simulations in total). These steps, carried out by adapting a forward selection process, make it possible to rank the most informative factors regarding the segments' condition.

In the first step, 8 simulations are performed. The UDB contains only one factor at a time. A randomly-chosen stock of segments (10% of the total length of the SVAS) is used to initially calibrate the STM (for each simulation and for each Monte Carlo try, a different sample is chosen randomly). The "age" factor with the largest η_4 (or $\mu_{30\%}$ c.f. section 4.3.4) is then chosen as the single most informative factor (table 4-10).

Table 4-10: simulation results for the first step

	Age	**Depth**	**Length**	**Material**	**Road class**	**Size**	**Gradient**	**Sewer type**
$\rho_{tar,4}$	13.4%	6.4%	6.4%	8.6%	7.6%	7.7%	9.7%	6.4%
η_4	2.0	1.0	1.0	1.3	1.1	1.2	1.5	1.0
$\mu_{30\%}$	38%	30%	30%	33%	31%	31%	34%	30%

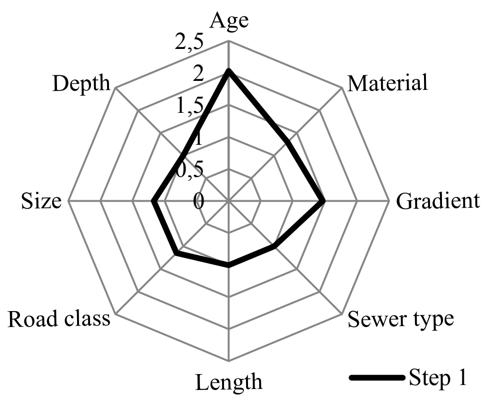
In the next step, 7 simulations should be performed as age has been already identified as the most informative factor. This time, the UDB will consist of age plus a different factor for each simulation. Similarly to step 1, by calibrating the STM using a randomly-chosen initial stock of segments whose length is equal to 10% of total length of the SVAS, the factor "material" is selected as it has the largest η_4 (or $\mu_{30\%}$).

This forward selection procedure will continue until all factors have been added to the UDB. This is done by including at each step all previously selected factors plus one of the remaining ones. Table 4-11 recapitulates the output for each step. The evolution of η_4 at all steps is also shown in figure 4-6.

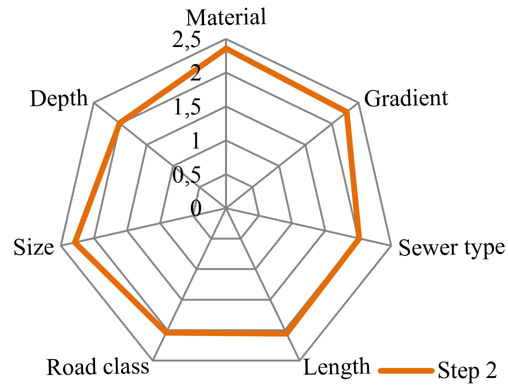
Table 4-11: output of different steps

Step	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
output	Age	+Material	+Gradient	+Sewer type	+Length	+Road class	+Size	+Depth
$\boldsymbol{\rho_{tar,4}}$	13.4%	15.6%	17.5%	17.9%	17.9%	17.9%	17.9%	17.9%
$\boldsymbol{\eta_4}$	2.0	2.4	2.6	2.7	2.7	2.7	2.7	2.7
$\boldsymbol{\mu_{30\%}}$	38%	41%	43.0%	43.5%	43.5%	43.5%	43.5%	43.5%

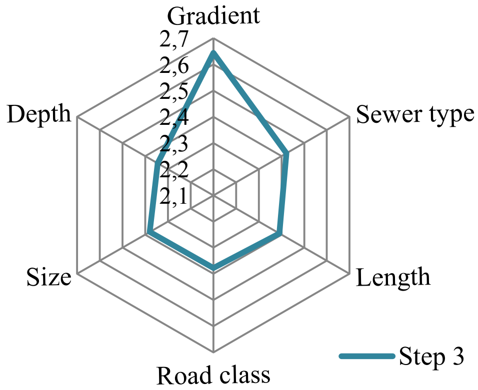
It should be noted that η for an inspection program based on a random principle (figure 4-2) is equal to 1. Consequently, in our case, having only age in the UDB doubles the inspection programs' efficiency. On the other hand, having only sewer type, length or depth in the UDB has no significant effect on inspection program efficiency compared to random-based inspection programs (table 4-10). In addition, where all factors are available within the UDB (full model), efficiency is 2.7 times better than random-based inspection programs. However, maximum efficiency is attained after selecting four most informative factors (table 4-11).



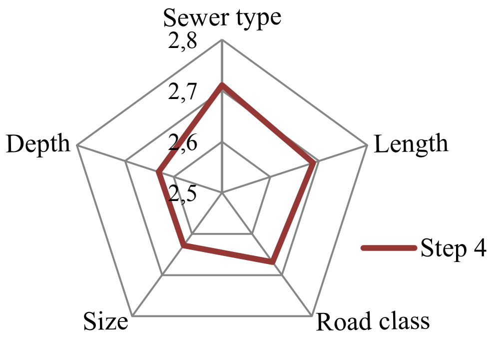
Step 1. Output: age. The UDB contains only one factor in each simulation.



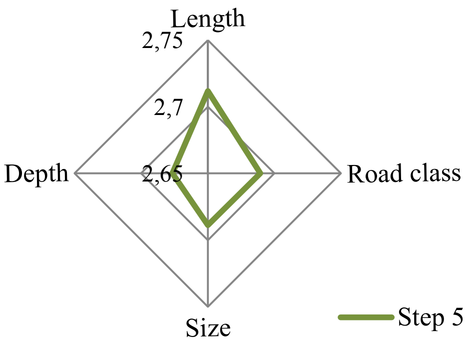
Step 2. Output: material. The UDB consists of age plus another factor in each simulation.



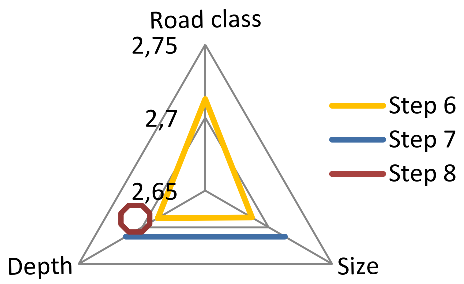
Step 3. Output: gradient. The UDB consists of age, material plus another factor in each simulation.



Step 4. Output: sewer type. The UDB consists of age, material, gradient plus another factor in each simulation.



Step 5. Output: Length. The UDB consists of age, material, gradient, sewer type plus another factor in each simulation.



Step 6-7-8. Output (respectively): Road class, size and depth.

Figure 4-6: The evolution of η_4 in different steps

4.5.1.2. Comparison of inspection programs efficiency and deviance dispersion

The ranking of the most informative factors (age followed by material, gradient, sewer type, length, road class, size and depth) found using the notion of inspection program efficiency (*a* *posteriori* analysis) is similar to that obtained using the backward selection method based on the deviance statistic (*a priori* analysis, presented in section 4.4.4.3).

Hilbe (2007) defines the deviance dispersion as the cumulative change in deviance divided by degree of freedom introduced into the model at each step. According to figure 4-7, for the first four variables a linear relation can be found between η_4 and change in deviance calculated using the backward selection method. However, η_4 does not change when introducing more than 4 factors into the UDB. This would be due to the existence of non-linear correlation(s) between these variables and the four most informative factors.

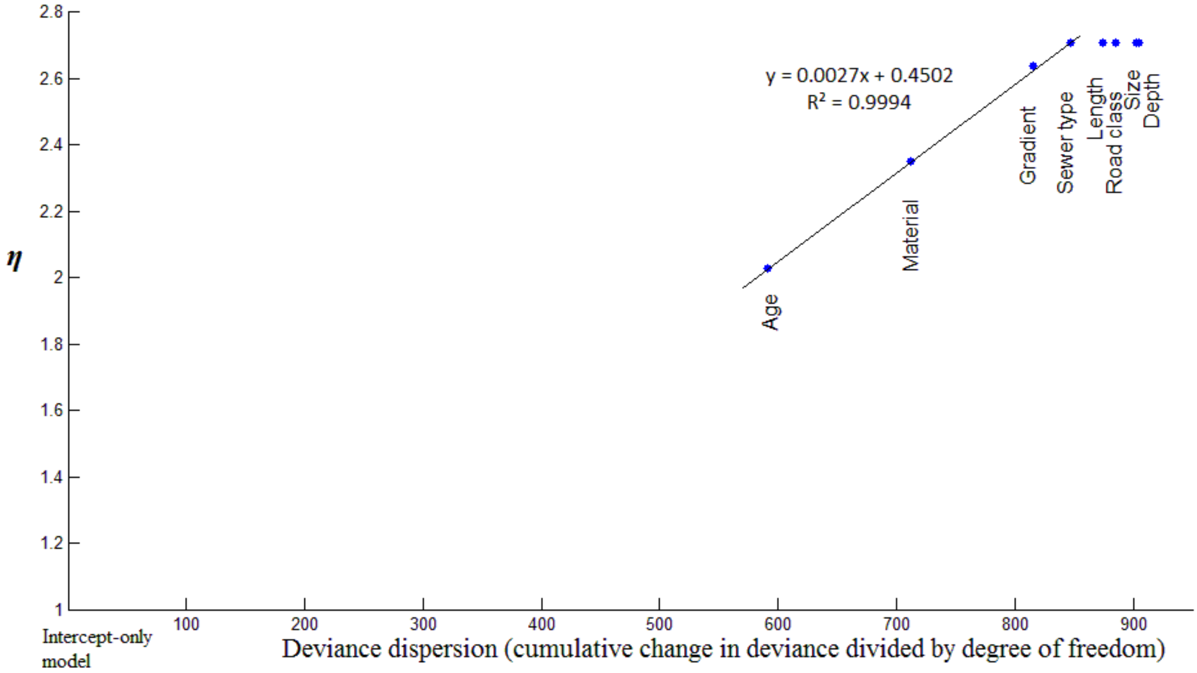


Figure 4-7: Comparison of two approaches of ranking the informativeness order of influential factors

Using a representative sample of the asset stock, containing several possible influential factors, the *a priori* analysis ranks the factors according to their informativeness and therefore according to the expected gain from acquiring data on a given factor for the rest of the asset stock. This result could then be used to plan the data acquisition programs if data is not available across the whole asset stock.

4.5.2. Influence of data imprecision and incompleteness: is imprecision preferable to incompleteness?

The aim of this section is to study influence of data imprecision and incompleteness by using different forms of "age" (as the most informative factor) in the UDB for the inspection programs. In other words, the main question is: "is it worth acquiring the exact age of segments or is a period of construction sufficiently informative to plan inspection programs?" In order to answer this question, six simulations were carried out. Age was available in different forms according to the simulation definition. All other factors were also available in the UDB for all cases:

(1) Age in the form of a scale variable: the exact construction year (age) of segments

(2) In the form of 2 age groups: periods with different management policies, war, construction technology and etc. Segments are divided into 2 categories according to their ages: ≤50 or >50 years old

(3) In form of 2 age groups: ≤80 or >80 years old

(4) In form of 4 age groups: ≤50; 50<age≤80; 80<age≤110 or >110

(5) In form of 7 urbanization periods ("districts")

(6) Without any information about age (data incompleteness)

The results of these simulations (table 4-12 and figure 4-8) show that:

(1) Depending on the adapted threshold used to divide the scalar data into 2 age groups, age in the form of 2 age groups can be significantly informative or not, in comparison with cases where no information is available about age within the UDB. We intentionally chose threshold values equal to 50 and 80 years old as the former value divides the asset into two groups with equal number of segments (50% of asset stock) as opposed to the latter value which assigns ¼ segments to category 1

and the remainder to category 2 (the data is more accurate with a threshold value equal to 80). It is noteworthy to mention that the categorization of scalar data would bring more information if the partitioning of different categories is based on different periods of management policies, remarkable changes in construction and installation technology, new pipeline fabrication technology and etc.

(2) Comparing the results of table 4-10 and 12, inspection program efficiency when the UDB contains only “age” and when all factors except “age” are available, is almost equal.

(3) “Age” as a scalar variable and “districts” in the UDB, compared to the case where no information is available on “age”, respectively improves inspection programs efficiency by a factor of 1.35 ($\eta 4$ = 2.7 compared to $\eta 4$ = 2.0) and 1.25 ($\eta 4$ = 2.5 compared to $\eta 4$ = 2.0). The fact that the notion of “district” could be used instead of age as a scalar variable can be generalized to all cases under the hypothesis that a sewer network is constructed at the same time. It is also obvious that acquiring information about districts where segments are laid is easier and less expensive than acquiring information on the precise age of segments. However, attention should be paid to the previously rehabilitated segments as they no longer have the same age as the district.

Table 4-12: the simulation results of imprecision and incompleteness of data within the UDB

Age represented in the UDB as:	**a scalar variable**	**2 age groups (≤50 or >50)**	**2 age groups (≤80 or >80)**	**4 age groups**	**Districts**	**Without any information on age**
$\rho_{tar,4}$	17.6%	13.1%	14.8%	17.0%	16.4%	13.0%
η_4	2.7	2.0	2.2	2.6	2.5	2.0
$\mu_{30\%}$	43.2%	37.6%	39.6%	42.5%	41.6%	37.5%

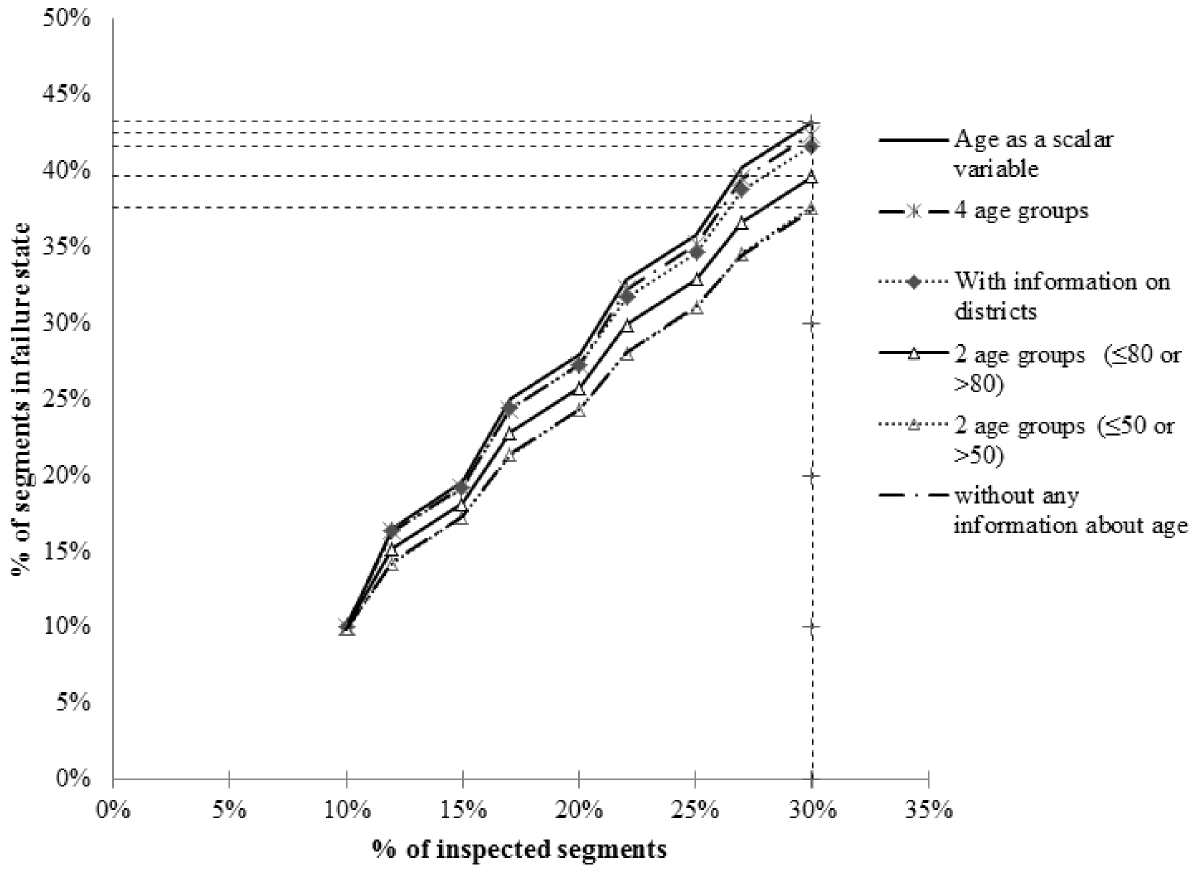


Figure 4-8: influence of imprecision and incompleteness of factor age on inspection programs

4.5.3. Influence of data incompleteness: Use of basic data and contribution of an auxiliary variable

In order to respond to the third group of questions evoked in table 4-2, we consider four utility databases corresponding to a four different knowledge of the same asset. For each database, the available influential factors are described in table 4-13.

Table 4-13: Available factors in each database

Factor available in the complete database	**Database 1: *basic data***	**Database 2: *basic data* + district**	**Database 3: *basic data* + precise age**	**Database 4: all factors available**
Precise age	N/A	N/A	+	+
Material	N/A	N/A	N/A	+
Gradient	+	+	+	+
Size	+	+	+	+
Sewer type	+	+	+	+
Length	+	+	+	+
Depth	+	+	+	+
Road class	N/A	N/A	N/A	+
District	N/A	+	N/A	N/A

For each database, a forward selection process (based on the framework described in section 4.3.1) is carried out in order to rank factors from the most to the least influential on condition grade. According to this approach, a factor is added to the model one at a time. At each step, each factor that is not already in the model is tested for inclusion in the model. At the end of each step, the most significant factor is then added to the model. Therefore, at the end of the first step, the model contains only the most significant factor. In the each following step, the most significant factor (from remaining factors) is added to the model. The steps continue until all factors are included in the model. The most significant factor is the factor obtaining the highest η_4 value (eq. 4). A randomly-chosen stock of segments (10% of the total length of the asset stock) is used to initially calibrate the STM (for each simulation and for each Monte Carlo try, a different sample is chosen randomly).

For example for the database 2, in the first step, 6 simulations are performed. For each simulation, we keep only one influential factor in the simulation's database. The factor "district" with the highest η_4 (or $\mu_{30\%}$ c.f. section 4.3.1 and table 4-14) is then chosen as the most single informative factor. In the next step, 5 simulations are performed as "district" has been already identified as the most informative factor. This time, the simulation's database will consist of "district" plus a different factor for each simulation. Similarly to step 1, by calibrating the STM using a randomly-chosen initial stock of segments whose length is equal to 10% of total length of the asset stock, the factor “gradient” is selected as it has the largest η_4 (or $\mu_{30\%}$).

Table 4-14: simulation results for the first step for database 2

	District	Depth	Length	Size	Gradient	Sewer type
$\rho_{tar,4}$	13.4%	6.4%	6.4%	7.7%	9.7%	6.4%
η_4	2.0	1.0	1.0	1.2	1.5	1.0
$\mu_{30\%}$	38%	30%	30%	31%	34%	30%

It should be noted that η_4 for an inspection program based on a random principle (choosing segments arbitrary) is equal to 1 (no amelioration on the inspection program efficiency as segments have been chosen randomly). Therefore, in first step for the database 2 having only depth, length or sewer type in the simulation’s database has no significant effect on inspection program efficiency compared to the random-based inspection program. On the other hand, having only district in the database doubles the inspection programs’ efficiency (correlated strongly to age).

This forward selection procedure will continue until all factors have been added to the simulation’s database. This is done by including at each step all previously selected factors plus one of the remaining ones. Table 4-15 recapitulates the output for each step for the database 2.

Table 4-15: output of different steps for database 2

Step	1	2	3	4	5	6
output	District	+Gradient	+Size	+Length	+Depth	+Sewer type
$\rho_{tar,4}$	13.4%	15.0%	15.8%	15.8%	15.8%	15.8%
η_4	2.0	2.3	2.4	2.4	2.4	2.4
$\mu_{30\%}$	38%	40%	41.1%	41.1%	41.1%	41.1%

Table 4-16 recapitulates the informativeness order of factors available for all databases (including the database 2). It also provides the η_4 for each simulation including the corresponding factor plus all previously chosen factors. For example, size is the third variable regarding the informativeness order of factors available in the database 2. In addition, adding size in the utility database which already contains district and gradient in third step increases η_4 from 2.3 to 2.4. Figure 4-9 shows the simulations results for these 4 utility databases.

Table 4-16: Simulation results, the ultimate efficiency (η_4) is underlined between parentheses

Factor available in the complete database	Informativeness order of single variable (and η_4)			
	database 1: *basic data*	database 2: *basic data* + district	database 3: *basic data* + precise age	database 4: all factors available
Precise age	N/A	N/A	1 (2.0)	1 (2.0)
Material	N/A	N/A	N/A	2 (2.4)
Gradient	1 (1.5)	2 (2.3)	2 (2.3)	3 (2.6)
Size	2 (1.8)	3 (2.4)	3 (2.4)	7 (2.7)
Sewer type	3 (1.8)	6 (2.4)	4 (2.5)	4 (2.7)
Length	5 (1.8)	4 (2.4)	5 (2.5)	5 (2.7)
Depth	4 (1.8)	5 (2.4)	6 (2.5)	8 (2.7)
Road class	N.A	N/A	N/A	6 (2.7)
District	N/A	1 (2.0)	N/A	N/A

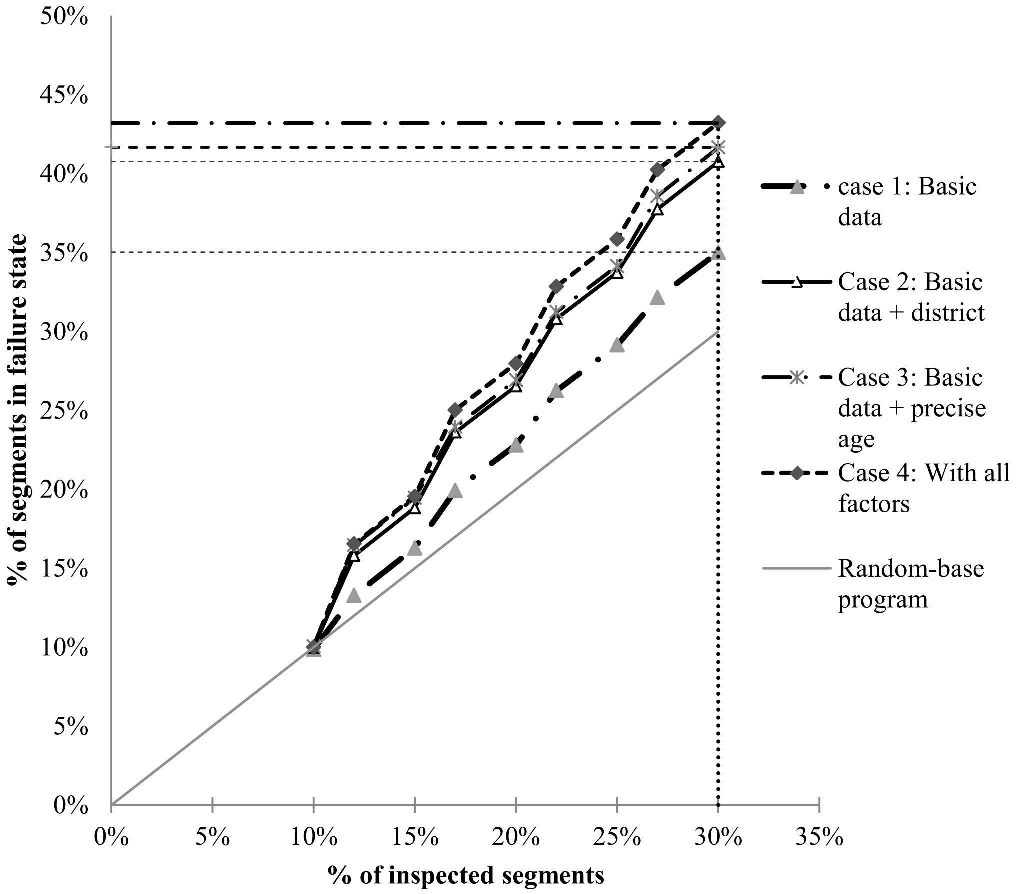


Figure 4-9: different simulations in order to study the influence of missing variable and auxiliary variable on the inspection programs

Following conclusions could be drawn from figure 4-9 and table 4-16:

(1) Having only basic data within the database nearly doubles the efficiency of inspection programs compared to a random-based one. It is important to note that in general these data are available in almost all utilities for hydraulic models.

(2) Adding an auxiliary variable does improve the efficiency of inspection programs. We showed that having district improves this efficiency about 30% compared to the case 1 where district is not available ($\eta 4$=2.4 compared to $\eta 4$=1.8 for case 1) as it is strongly correlated with age.

(3) Having only district offers better efficiency than having only basic data ($\eta 4$=2.0 instead of $\eta 4$=1.8) as we have shown earlier, age is the most informative variables in terms of inspection program efficiency and the factor "district" is strongly correlated to this factor.

(4) Sometimes, having the precise information about assets is not necessary. Having district almost has the same effect as precise age on the efficiency of inspection programs ($\eta 4$=2.4 for the database 2 compared to $\eta 4$=2.5 for the database 3).

(5) The efficiency of the database where all factors are available is only 13% better than for the database where it only consists of basic data plus district ($\eta 4$=2.7 for case 4 and 2.4 for case 2). It is noteworthy to mention that acquiring information about segments precise age and material could be very expensive or almost impossible in many cases. Therefore, depending on the utilization purpose of data and due to a lack of funds, this information could be replaced by other factors easily accessible if they are correlated with the missing ones.

(6) In all databases, we achieve the maximum efficiency by having only first four most important variables.

4.5.4. Influence of data uncertainty: Is it worth to accept a degree of uncertainty within data instead of not having them

The goal of this section is to study the influence of having uncertainty within the "age" factor (the most informative variable for the studied asset stock regarding the efficiency of inspection programs). In other words, the main questions are: "is it worth to accept a degree of uncertainty within data instead of not having them?" and "how does the uncertainty affect the efficiency of inspection programs?"

4.5.4.1. Level of uncertainty

In order to introduce uncertainty into the asset stock and to form the utility databases for the scenarios defined in section 4.5.4.2, we consider the case where information about "age" is available in form of 4 age groups: ≤50; 50<age≤80; 80<age≤110 or >110. Afterwards, we intentionally misplace a certain percentage of segments into the categories where they do not belong. In total, we consider 3 degrees of uncertainty as follows:

- 10% of segments in each age group are misplaced;
- 20% of segments in each age group are misplaced;
- 30% of segments in each age group are misplaced;

Assume that $u \in \{0.1, 0.2, 0.3\}$ (called *uncertainty ratio*: each u corresponds to a specified case). For each segment, a random number between 0 and 1 is generated. If this number is smaller than u, the segment will be misplaced into another age group. Then, if the segment should be misplaced, by considering the real age group of the segment, another random number between 0 and 1 is generated and is compared with the cumulative misplacement probabilities given in table 4-17. Afterwards, the misplaced age group for the segment is assigned. For example, assume that u=0.2 (case 2) and the real condition grade for segment j is G3. If the first generated random number is smaller than 0.2 (for example 0.14), then the segment should be assigned into another age group. Assume also that the second random number is equal to 0.42. Then, considering table 4-9, the new (misplaced) age group is G2.

Table 4-17: Cumulative probability of being in a given misplaced age group for a segment selected to be misplaced in first step

		misplaced age group			
		G1	G2	G3	G4
Real age group	G1	---	0.7	0.9	1
	G2	0.4	---	0.8	1
	G3	0.2	0.6	---	1
	G4	0.1	0.3	1	---

Figure 4-10 shows the evolution of segment distributions between age groups for these cases plus the case where no uncertainty is available within the factor "age". It is observed that by applying this procedure, in general the asset stock tends to get older.

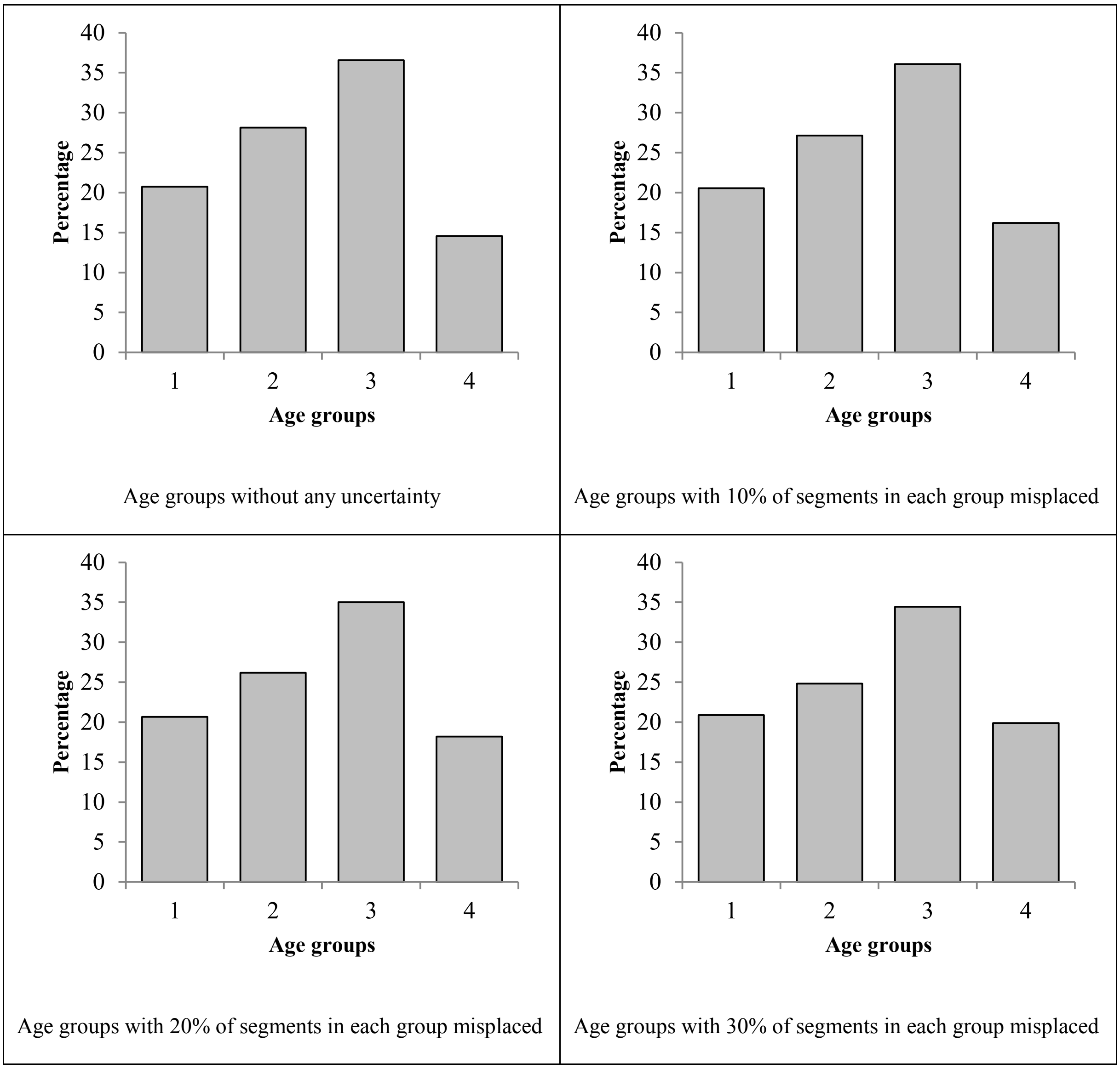


Figure 4-10: Evolution of misplaced segments percentage in each age group

4.5.4.2. Utility databases and defined scenarios

Once uncertainty is inserted within the age factor, we proceed to create the utility databases in order to respond to the evoked questions in section 4.3.2. Therefore, in total 10 scenarios are defined (table 4-18). Considering the presence of material and road class factors within the utility database, these scenarios could be classified into two main groups:

(1) Scenarios (databases) in which these factors are not available. They are then called “basic data related scenarios” as the utility database in these cases consists of sewer size, sewer type, gradient, depth and length of segments plus possible presence of the factor “age” in a defined form.

(2) Scenarios (databases) in which all factors are available in their precise formats except the factor “age” which could be available in a defined form.

Table 4-18 illustrates the availability of each factor within the defined scenarios. It is important to note that S10 is the scenario which represents the incompleteness (of factor age) and scenarios S6 to S9 represent the uncertainty (within factor age as well).

Table 4-18: the availability of each factor within defined scenarios, age is always presented in form of 4 age groups.

Factor	*Basic data* related scenarios					Complete database related scenarios				
	S1: without uncertainty	S2: with 10%	S3: with 20%	S4: with 30%	S5: with no information about age	S6: without uncertainty	S7: with 10%	S8: with 20%	S9: with 30%	S10: with no information about age
Age with no uncertainty	+	N/A	N/A	N/A	N/A	+	N/A	N/A	N/A	N/A
Age with 10% uncertainty	N/A	+	N/A	N/A	N/A	N/A	+	N/A	N/A	N/A
Age with 20% uncertainty	N/A	N/A	+	N/A	N/A	N/A	N/A	+	N/A	N/A
Age with 30% uncertainty	N/A	N/A	N/A	+	N/A	N/A	N/A	N/A	+	N/A
Material	N/A	N/A	N/A	N/A	N/A	+	+	+	+	+
Gradient	+	+	+	+	+	+	+	+	+	+
Size	+	+	+	+	+	+	+	+	+	+
Sewer type	+	+	+	+	+	+	+	+	+	+
Length	+	+	+	+	+	+	+	+	+	+
Depth	+	+	+	+	+	+	+	+	+	+
Road class	N/A	N/A	N/A	N/A	N/A	+	+	+	+	+

4.5.4.3. Simulation results

For each scenario introduced above, the simulation framework described in section 4.3.1 has been carried out. The results are provided in table 4-19 and figure 4-11. Following conclusions could be drawn:

(1) Having age in an imprecise form of 4 age groups is nearly as informative as having age in a scale form (case 4 of previous section) regarding the influence of this specific data on inspection program efficiency ($\eta4$= 2.6 where age is in 4 age groups form and $\eta4$=2.7 where precise age is available).

(2) It is shown that having uncertain data is more informative than not having this data within the utility database.

(3) Depending on the uncertainty ratio, the uncertain data could considerably ameliorate the efficiency of inspection programs. However, it should be noted that for the extreme case, having 30% uncertainty on a segment's age group means that by a

chance of 1 to 3, we have failed to correctly assess this information. In practice, by considering the chosen thresholds for age groups and other factors such as construction time of surface development, type of material and …, this is unlikely to have this huge amount of uncertainty.

(4) In our case, by adding material and road class into the utility database, only an improvement of 15% occurs in the efficiency of inspection programs (η4=2.3 for S1 and 2.6 for S6).

(5) An inspection program with only the basic data plus imprecise uncertain information about age (in form of 4 age groups with 30% of uncertainty) is twice more efficient than a random-base program (η_4=2 for the former and η_4=1 for the latter).

Table 4-19: Defined scenarios simulation results

	Basic data related scenarios					Complete database related scenarios				
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
$\rho_{tar,4}$	15.0%	14.4%	14.1%	13.4%	12.1%	17.0%	16.7%	16.0%	15.1%	13.0%
η_4	2.3	2.2	2.1	2.0	1.8	2.6	2.5	2.4	2.3	2.0
$\mu_{30\%}$	40.0%	39.8%	39.5%	37.9%	35.0%	42.5%	41.8%	41.3%	40.1%	37.5%

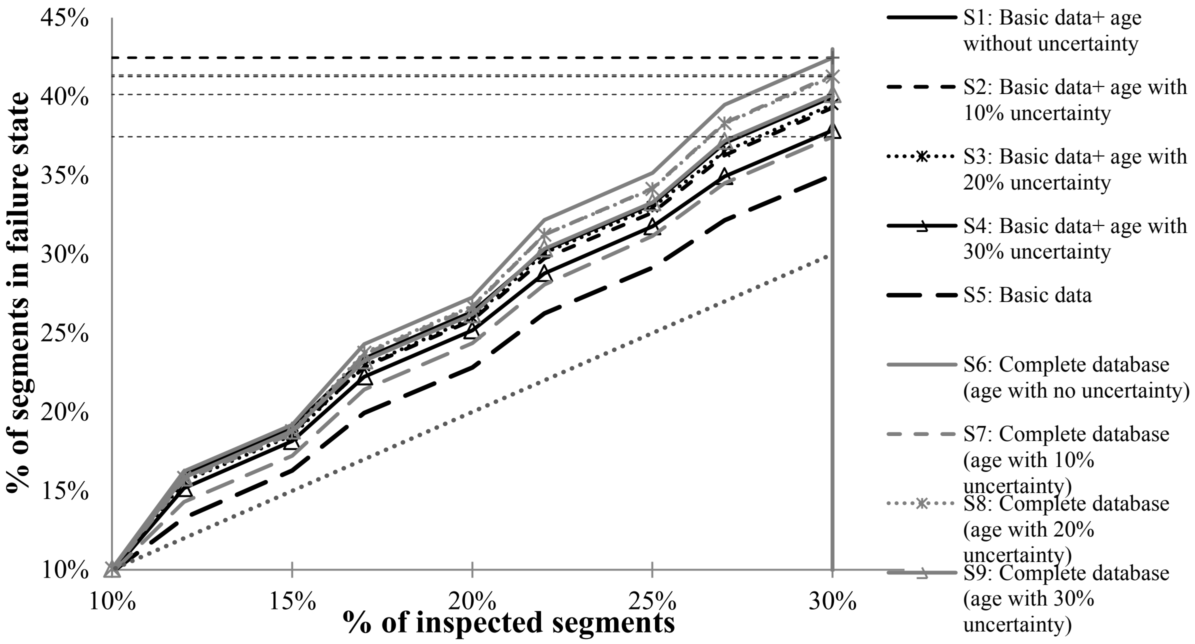


Figure 4-11: different simulations in order to study the influence of uncertainty on the inspection programs

In this section we tackled the problem of uncertainty within "age" as an influential factor. However, the uncertainty is also present within the dependent variable which is the condition grade of segments. Dirksen *et al.,* (2011) mention the main reasons and major types of uncertainty from visual inspection towards the condition grade of a segment. However, the influence of uncertainty within condition grade of segments on the efficiency of inspection programs needs further investigations.

4.5.5. Recommendations for applying the proposed methodology on real cases

Assume that a utility possesses a representative sample of its asset stock for which possible influential factors plus segments condition grade are available. The application of our methodology on real cases, in order to establish the list of most informative variables involves the implementation of all of the following steps (as it is described progressively along the paper):

(1) Analysing the available data to test the significance of each factor;

(2) Generating combinations of variables in order to study the possible interactions between influential factors;

(3) Applying the backward selection method based on the deviance statistic (likelihood ratio test), with single variables and combinations of variables generated in second step;

(4) Selecting most informative variables.

In our case, by considering only the single variables, we showed that the factor "age" is the single most informative factor regarding the inspection program efficiency followed by "material", "gradient" and "sewer type". It is also observed that adding the factors "length", "road class", "size" and "depth" does not cause any differences in the inspection program efficiency.

On the other hand, by considering the difficulty of acquiring specific data about assets, we showed that in many cases, using imprecise data such as “district” or “age” groups could largely replace the precise age of segments. Therefore, it is heavily recommended to gather information about most informative factors in any possible format if the precise data is not available.

4.6. Conclusion

A vital component of proactive programs is the assessment of sewer segments which relies mostly on visual inspection.

The process for defining visual inspection priorities can take into account the predicted condition together with environmental characteristics to consider failure consequences.

The efficiency of sewer inspection program improves if the inspection surveys target segments in failure state. Therefore, the accurate prediction of current and future condition of sewers is crucial for effective decision-making. This prediction can be made using deterioration models.

We proposed a systematic approach using a deterioration model (logistic regression) to improve sewer inspection program efficiency. The influence, on the inspection program efficiency of the data available within a utility was also studied by adapting this systematic approach for use as a framework for numerical simulations.

In order to study the influence on the inspection programs of the data available within a utility, a semi-virtual asset stock was created and the corresponding database was then degraded to introduce imprecision, incompleteness and uncertainty into the utility database. The main questions concerning these aspects were:

(1) Which factor is most informative amongst the available ones? And how to establish the list of most informative single factors?

(2) Is it preferable to have imprecision instead of incompleteness within the utility database?

(3) How can the data most probably available within a utility be used to define an effective inspection program? For example, are data collected for hydraulic models of network such as segment's diameter, depth, gradient, sewer type and length, called basic data, sufficient and relevant enough for inspection programs?

(4) Can we use an auxiliary variable in order to compensate effects of missing data?

(5) Is it worth to accept a degree of uncertainty within data instead of not having them?

Two different approaches were proposed to answer the first question: 1- a backward selection method using the deviance statistic (likelihood ration test) and adapting a nested subset process and 2- a forward selection method applying the developed systematic approach for inspection programs based on an indicator to gauge inspection program efficiency. The results for two approaches were identical. This makes it possible to use the first approach (based on deviance statistic) on a representative sample of an asset stock to establish the list of the most informative factors and also to plan the data acquisition programs.

In order to answer the second question, we considered six different databases with imprecision and incompleteness compared to the complete database with the factor "age" being the most informative of all the factors available in this case. Numerical simulations yielded interesting results which contribute significantly to the small number of studies previously carried out in this area. This study would suggest that imprecision in the database is preferable to having a database with no information on one specific factor. We also showed that the notion of "district" could be used instead of precise information on segment age under the hypothesis that a sewer network is constructed at the same time as the surface development above it.

For other questions, the results show that:

(1) Easily accessible and available data (basic data) could improve considerably the efficiency of inspection programs;

(2) The use of an auxiliary variable such as geographical position could compensate the effect of missing data and it does improve the efficiency of inspection programs.

(3) Having data even with huge amount of uncertainty is preferable to having incompleteness within the utility database.

However, following points need further investigations:

- The influence of initial stock of segments which is used in order to calibrate the deterioration model. This issue is partially addressed in chapter 6.
- We showed that the backward selection method based on the deviance statistic produces similar results as inspection programs for the whole asset stock. In other words, this approach was applied on the whole asset stock. However, in the reality, this is not the case. Therefore, we should study outcomes of the application of this method on a representative sample of our asset stock and compare the results to see if similar results are again produced.
- An economic analysis of inspection programs defined in order to gauge benefits that a utility could make by defining these programs;
- In general, other deterioration models may have better efficiency or not. There is no solid recommendation concerning the choice of deterioration model within the scientific literature. Hence, this choice could be the objective of further investigations.

All these results will be useful for demonstrating the gain obtained from acquiring data and implementing models on sewer asset management to wastewater utilities.

4.7. References

1. Agresti, A., 2002. “Categorical data analysis”, 2nd Ed. Hoboken, NJ: Wiley.
2. Ahmadi M., Cherqui F., De Massiac J.C., Le Gauffre P., (2013a). “Influence of available data on efficiency of sewer inspection programs”. *Urban water journal*. (DOI:10.1080/1573062X.2013.831910).
3. Ahmadi M., Cherqui F., De Massiac J.C., Le Gauffre P., (2013b). “Benefits of using basic, imprecise or uncertain data for elaborating sewer inspection programs”. *Structure and Infrastructure Engineering*. (Accepted)
4. Ana, E. and Bauwens, W., (2010). “Modeling the structural deterioration of urban drainage pipes: the state-of-the-art in statistical methods”. *Urban Water Journal*, 7(1), 47 – 59.
5. Ariaratnam, S. T., El-Assaly, A. and Yang, Y., (2001). “Assessment of infrastructure inspection needs using logistic models”. *Journal of Infrastructure System*., 74, 160–165.
6. Arthur, S. and Burkhard, R., (2009). “Prioritizing Sewerage Maintenance using Inferred Sewer Age - A Case Study for Edinburgh”. *Water science and technology*, 61(9), 2417-2424.
7. Baik, H. S., Jeong, H. S. D. and Abraham, D. M., (2006). “Estimating transition probabilities in Markov chain-based deterioration models for management of wastewater systems”. *Journal of Water Resource Planning and Management*. 132(1), 15–24.
8. Bailey, T. and Jain, A., (1978). “A note on distance-weighted k-nearest neighbor rules”. *IEEE Trans. Systems, Man, Cybernetics*, Vol. 8, pp. 311-313.
9. Breysse, D., Vasconcelos, E. and Schoefs, F., (2007), “Management strategies and improvement of performances of sewer networks”, *Computer Aided Civil and Infrastructure Engineering*, 22(7), 462–77.
10. Cramer, D., (1994). “Introducing statistics for social research: step-by-step calculations and computer techniques using SPSS”. London; New York: Routledge
11. Davies, J. P., Clarke, B. A., Whiter, J. T. and Cunningham, R. J., (2001a). “Factors influencing the structural deterioration and collapse of rigid sewer pipes”. *Urban Water Journal*, 3(1–2), 73–89.
12. Davies, J. P., Clarke, B. A., Whiter, J. T., Cunningham, R. J. and Leidi, A., (2001b). “The structural condition of rigid sewer pipes: A statistical investigation”. *Urban Water Journal*, 3(4), 277–286.
13. Fix, E. and Hodges, J.L., (1951). “Discriminatory analysis, nonparametric discrimination: Consistency properties. Technical Report 4”, USAF School of Aviation Medicine, Randolph Field, Texas.
14. Gemora, I., (2003). “Pipeline assessment and certification program”. *Proc. Int. Conf. on Pipeline Engineering and Construction*, ASCE, Reston, VA, 822–825.
15. Guyon, I. and Elisseeff, A., (2003). “An Introduction to Variable and Feature Selection”. *Journal of Machine Learning Research*, 3, 1157-1182.

16. Hilbe, J., (2007). “Negative Binomial Regression”. Cambridge University Press.
17. Kleiner, Y.; Rajani, B.B, Krys, D. (2013). “Performance of ductile iron pipes: characterization of external corrosion patterns”, *ASCE Journal of Infrastructure Systems*, 19(1), pp. 108-119.
18. Koo, D.-H. and Ariaratnam, S. T., (2006). “Innovative Method for Assessment of Underground Sewer Pipe Condition”, *Automation in Construction*, 15, 479-488.
19. Le Gauffre, P., Joannis, C., Vasconcelos, E., Breysse, D., Gibello, C., Desmulliez, J.J., (2007). “Performance Indicators and Multicriteria Decision Support for Sewer Asset Management”. *Journal of Infrastructure Systems*, ASCE, 13(2), 105-114.
20. McCullagh, P. and Nelder, J.A., (1989). “Generalized Linear Models”, second ed. Chapman & Hall/CRC, USA.
21. Möderl, M., Sitzenfrei, R., Fetz, T. and Rauch, W., (2011). “Systematic generation of virtual networks for water supply”. *Water Resources Research*. 47.
22. Montgomery, D.C., Runger, G.C., and Hubele, N.F., (2004), “Engineering Statistics”, JohnWiley & Sons, Massachusetts.
23. Salman, B., (2010). “Infrastructure management and deterioration risk assessment of wastewater collection systems”. Thesis (PhD), Univ. of Cincinnati, OH.
24. Salman, B., and Salem, O., (2012). “Modeling failure of wastewater collection lines using various section-level regression models”, *Journal of Infrastructure Systems*, ASCE, 18(2), 146-154.
25. Scheidegger, A., Hug, T., Rieckermann, J. and Maurer, M., (2011). “Network condition simulator for benchmarking sewer deterioration models”. *Water Research*, 45(16), 4983-4994.
26. Selvin, S., (1994). “Practical biostatistical methods”. Belmont, Calif: Wadsworth Publishing Co, 378–380.
27. Tran, D.H., Perera, B.J.C. and Ng, A.W.M., (2009). “Comparison of structural deterioration models for stormwater drainage pipes”. *Comput. Aided Civ. Infrastruct. Eng*. 24 (2), 145–156.
28. Urich, C., Sitzenfrei, R., Möderl, M. and Rauch, W., (2010). “An agent-based approach for generating virtual sewer systems”. *Water Science and Technology*, 62(5), 1090-1097.
29. Yang, M. D. and Su, T. C., (2007). “An optimization model of sewage rehabilitation”. *Journal of the Chinese Institute of Engineers*, 30(4), 651–659.
30. Zhao, J.Q., (1998). “Trunk sewers in Canada”. In: 1998 APWA International Public Works Congress Seminar Series, American Public Works Association. Las Vegas, Sept. 14–17.

Chapter V: Assessment of a segment by an inspection

5.1. Overview

This chapter is extracted from a paper entitled "Condition grading for dysfunction indicators in sewer asset management" published in Structure and Infrastructure Engineering (Ahmadi et al., 2013).

Infrastructure asset management relies on the establishment of inspection and rehabilitation priorities. During the last ten years an increasing number of studies have been dedicated to rehabilitation needs and priorities, combining an asset's performance and physical condition. International Water Association (IWA) working groups proposed performance indicators for water (Alegre *et al.,* 2000) and for sewer systems (Matos *et al.,* 2003). In Europe, several large R&D projects were dedicated to defining rehabilitation priorities using both condition assessment and performance indicators: CARE-W (Saegrov, 2005) for water supply networks, CARE-S (Saegrov, 2006) for wastewater networks, APUSS (Ellis & Bertrand-Krajewski, 2006) related to infiltration and exfiltration of urban sewer systems. Several operational tools were also developed to define inspection needs and priorities for sewer pipes, such as SCRAPS (Hahn *et al.,* 2002; Marlow *et al.,* 2007, 2008; Merrill *et al.,* 2004).

Visual inspection and more specifically closed circuit television inspection (CCTV) is now one of the main methods used to assess the condition of sewer networks and is widely used (Knolmar and Szabo, 2003). The NRC-CNRC in Canada produced MIIP reports (Municipal infrastructure investment planning project) to separately assess structural and operational defects (Vanier *et al.,* 2004, 2006, 2006b). In France, the national RERAU methodology (Le Gauffre *et al.,* 2004; 2007) proposes a protocol to assess 10 dysfunction indicators from defect observation (table 5-1). We explained the defined dysfunction indicators in chapter 2, but they are repeated here as we are going to use them all along this chapter.

Table 5-1. Dysfunction and impact indicators assessed from visual inspection reports

Dysfunctions	Definition	Impacts	Definition
ABR	Ongoing degradation from abrasion	DAB	Damage to buildings (including infiltration of water into basements)
BLO	Blockage	NUH	Nuisance of a hydraulic nature (service interruption, flooding…)
COL	Risk of collapse	OCP	Treatment plant operating surplus costs
COR	Degradation due to corrosion	OCS	Network operation surplus costs (including the cost of equipment shortened lifetime)
CSO	Excessive spillage	POG	Pollution of ground and groundwater
EXF	Exfiltration (seepage loss)	POL	Pollution of surface water resources, due to overflows, spillage or disturbance to treatment system process
FLO	Flooding	SLC	Shortened lifetime cost, along with the surplus cost associated with remedial actions
HYD	Decrease in hydraulic capacity	TRA	Traffic disruption (including for the needs of sewer operation: Cleansing …)
INF	Infiltration		
ROO	Ongoing degradation from root intrusion		
SAN	Sand silting		
SPD	Destabilization of ground-pipe system		

5.2. Scope and objective of this chapter

Although visual inspection is widely used, many sources of uncertainty exist within the process of constructing rehabilitation criteria from the results of investigations. Four major types of uncertainties related to visual inspection data can be remarked as follows:

(1) Inspecting a segment and saving the existing defects by predefined codes by CCTV operator. Dirksen et al., (2011) illustrates how uncertainty is introduced into inspection reports by CCTV operators.

(2) Converting each code into a score can depend on the tables used for the conversion (discussed in section 5.3.2).

(3) Assigning a condition grade to a segment from its score (issue addressed in this chapter).

(4) The fourth type of uncertainty is related to the table used to aggregate an indicator derived from visual inspection with another indicator (for assessing a rehabilitation criterion). Crossing indicator X (grade G2) with indicator Y (grade G3) to assess indicator Z may lead to two possible crisp aggregation results: G2 by an optimistic and G3 by a pessimistic view. Which option should be chosen? This type of issue is addressed in Le Gauffre and Cherqui (2009), who compare crisp and fuzzy operators.

In section 5.3, we introduce a method to quantify the observed defects on a segment. We then justify the use of a single score as a representation of the global condition of a sewer segment.

Section 5.4 proposes a method to improve the process of assigning a condition grade to a segment.

In section 5.5, this method is applied on 150km of the Greater Lyon asset stock. A sensitivity analysis of parameters used in the proposed method, described in section 5.4, is also carried out for a better understanding of their impacts on the evaluation of an asset stock.

5.3. Translation of inspection results into a condition grade

Condition assessment of sewer segments is an important component of asset management and relies mostly on visual inspections (Ana and Bauwens, 2010; Knolmar and Szabo, 2003; Marlow *et al.,* 2007; Rahman and Vanier, 2004). Closed Circuit televised inspection (CCTV) provides an inventory of observations (generally defects) for the assessment of the internal condition of sewers. Condition assessment is made by professionals, either by subjective grading (i.e. expert's opinion based on experience), or using a qualitative scale ("distress-based evaluation" according to Rahman and Vanier, 2004).

5.3.1. Existing protocols to assess condition grade

Focusing on distress-based evaluation, the condition grade of a sewer segment is assigned by comparing a defects inventory (defect coding and deduct values) with a subjective scale of numerical values. The defects and qualitative scale are converted to identify structural problems (degradation, risk of collapse, etc.) or operational problems (blockage, infiltration, flooding, etc.). The most well-known protocols, as explained in chapter 3, are WRC's (Rahman and Vanier, 2004) "Sewerage Rehabilitation Manual" (Water Research Centre), NRC's "Guidelines for Condition Assessment and Rehabilitation for Large Sewers" (National Research Council of Canada) and protocols based on WRC, such as Winnipeg's "Sewer Management Study", Edmonton's "Standardized Sewer Condition Rating System Report", NAAPI's "Manual for Sewer Condition Classification" (North American Association of Pipeline Inspectors) and NASSCO's "Pipeline Assessment Certification Program" (National Association of Sewer Service Companies). These protocols use one or several of the following formula:

Mean or density or single score= sum of deducted values / length of segment (1)

Peak score= Maximum deducted value (2)

Total score= sum of deducted values (3)

The mean score or density score corresponds to the global condition of the sewer segment, the peak score enables the most serious defect to be identified and finally the total score represents the number and seriousness of defects on a sewer segment (qualitative and quantitative). The mean score and peak score may be used to compare segments of different lengths; however this is not the case for the total score. The relevance of a single score is justified in section 5.3.3.

On the other hand, each aforementioned protocol defines a condition grading range (Table 5-2). The lowest value (0 or 1) and highest values always correspond respectively to the best condition and the worst condition of a sewer segment.

Table 5-2. Condition grading range of different protocols, adapted from Rahman and Vanier, 2004.

Protocols	WRc, Edmonton, Winnipeg, NAAPI, NASSCO	NRC	RERAU
Condition range	1 – 5	0 – 6	1 – 4

However, a bigger range does not necessarily lead to greater precision in maintenance prioritization, because the context (the sewer segment's environment) should also be taken into consideration.

5.3.2. Defect quantification according to the RERAU methodology

For assessing the condition grade of a segment, the visual observations, obtained by CCTV inspections, are converted into sequences of codes. The coding system can be a general standard (for example, the European standard EN 13508-2), a national standard (for example AGHTM 1999 in France, ATV 1999 in Germany) or a utility's self-defined norm. The observation codes can then be quantified in order to obtain a distribution of scores on each sewer segment.

The RERAU methodology proposes a scoring system that *"has a possibility to process pinpoint defects as well as defects distributed all over the length of a segment, while taking into account both their levels of seriousness and their extents"* (Le Gauffre *et al.,* 2007).

For each dysfunction, each coded defect i is translated into an elementary score N_i by multiplying the level of seriousness of this defect by its extent. Extent is the length (L_i) of a continuous defect (like a longitudinal crack) or a predetermined value b attributed to pinpoint defects (like an intruding connection). The degree of seriousness of each defect depends on a single parameter a and four levels of gravity respectively representing minor, medium, serious and very serious defects: 1, *a*, a^2, a^3. Thus, each sewer segment is characterized with a set of scores N_i :

$$N_i = a^n \times b(or\ L_i), \text{with } n = 0, 1, 2\ or\ 3 \text{ and } a = 2, 3\ or\ 4 \quad (4)$$

Table 5-3 presents an extract from the evaluation table of indicator INF4 (infiltration assessed from CCTV inspection in accordance with EN13508-2).

Table 5-3. Assessment of INF4: defects contributing to infiltration (Le Gauffre *et al.,* 2004).

Observation Oi	Code Ci	1	a	a^2	a^3	← Gravity Extent ↓
Deformation	BAA		BAA			b
Fissure	BAB	BAB B		BAB C		Li
Break/collapse	BAC			BAC A	BAC B/C	b
Missing mortar	BAE		BAE			b
Defective connection	BAH			BAH B/C/D		b
…	…					

The sum of these scores is then divided by the sewer length and is assigned as the mean or density score of the segment (like WRc and NRC protocols). This score seems to be the most appropriate single score (Ibrahim *et al.,* 2007).

$$D = \frac{\sum_i N_i}{length\ of\ segment} \quad (5)$$

5.3.3. Relationship between a segment's score and its length

A study (Werey *et al.,* 2006) shows a good correlation between the total score of a sewer segment ($T = \sum_i N_i$) and its length (Figure 5-1).

In Figure 5-1, each point represents an average of around 300 segments for a given range of sewer length. Other distributions with different number of segments per point were also tried and results were identical. As shown, there is a good linear relationship between total score ($T = \sum_i N_i$) and segment length (with a coefficient of determination of 0.97) concerning sewer segments under 55m. The strong linear relationship ($T = 2.45L^{0.99} \cong 2.45L$) between total score and segment length confirms the ability of the single score ($D = \frac{T}{L}$) to characterize a segment. D is consequently independent from the length of segment. Further investigations are required to characterize segments of over 55m (Ibrahim *et al.,* 2007).

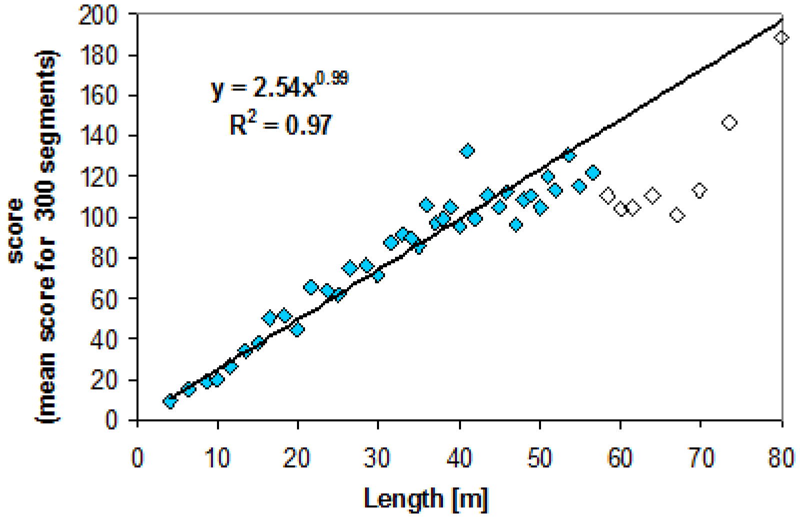


Figure 5-1. Score of the infiltration dysfunction indicator concerning more than 15 200 sewer segments from the department of Bas-Rhin, France (Werey *et al.,* 2006).

5.3.4. Scale of numerical variables

According to RERAU methodology, the single score (D) of a dysfunction indicator is compared to three thresholds (S1, S2 and S3) in order to define four condition grades:

- Grade 1 (G1: no or few detected defects) if D <= S1; (6)
- Grade 2 (G2: low gravity situation, segment to be monitored) if S1 < D <= S2;
- Grade 3 (G3: high gravity situation, intervention to be prioritized) if S2 < D <= S3;
- Grade 4 (G4: unacceptable situation in any context, needing action) if D > S3.

While most protocols (NRC, WRc …) assign an indicator to a grade by comparing its mean score with a subjective scale of numerical values (a set of pre-calculated thresholds), the RERAU methodology proposes a method to determine thresholds for each specific asset stock in order to take into account the specific features of each utility.

In the next section, we describe this method, which aims to optimize thresholds for each utility. Optimizing means minimizing a calibration criterion taking into account the specific features of each utility, such as asset stock general condition and local expectations regarding the condition of the assets (explained and detailed in section 5.4).

5.4. Assigning a segment to a condition grade

Figure 5-2 presents a comparison of two existing processes for assigning a segment to a condition grade: a) Direct assessment by engineers (expert opinion taken as reference); b) Comparison of the single score (D) with thresholds (equation 6).

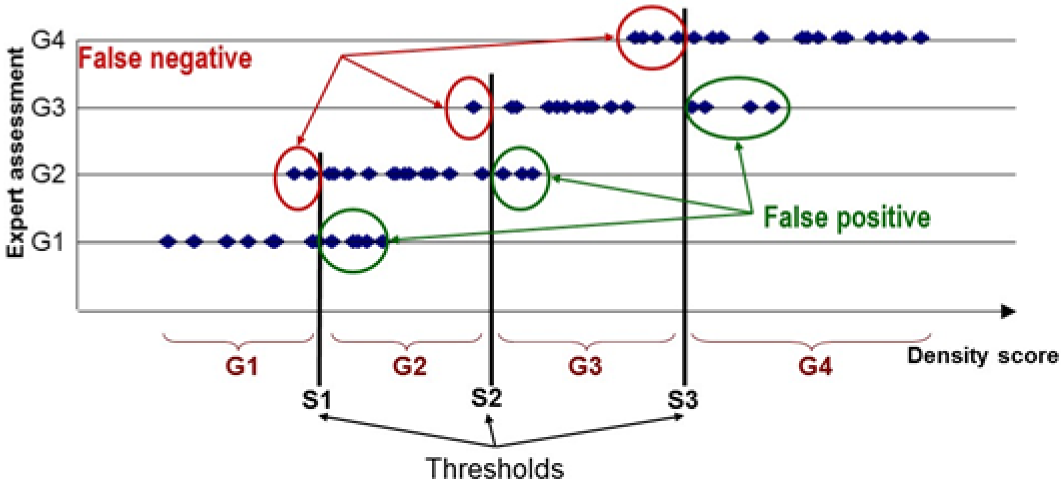


Figure 5-2. Expert opinions versus scores of a dysfunction indicator, for a given set of thresholds

In the next section, we explain the notion of false negative and false positive assignment error according to the second process (comparison of D with threshold values). The calibration criterion is then introduced in section 5.4.2. This criterion, based on a set of segments evaluated by experts (process a), is used to adapt the threshold values to the specific features of each utility. Details of calibration criterion requirements and processes are discussed in sections 5.4.3 to 5.4.5.

5.4.1. Control of assignment error

By considering a set of thresholds (Figure 5-2), two types of error can be highlighted. A false positive (FP) error is a segment wrongly assigned to a higher grade (worse condition). Conversely, a false negative error (FN) is a sewer segment assigned to a lower grade than that given by experts. For example, among all the sewer segments assessed as being in grade 3 by experts (considered as reference), some segments are estimated as lower grades by equation 6 (false negative: grade 1 or 2) and some segments are estimated as higher grades (false positive: grade 4).

Consequently, the optimum set of thresholds needs to be found and this means minimizing assignment errors by taking into account whether it is better to overestimate the condition grade of a segment rather than to underestimate its condition. The proposed calibration procedure is detailed in the next sections.

5.4.2. Calibration criterion

The threshold calibration process aims to minimize the consequences of the assignment errors mentioned in the previous section (false positive and false negative errors). This process allows the utility manager to decide the importance of each type of error. It also considers the characteristics of the studied asset stock. A calibration criterion *CC* is proposed for determining cut-off values (S1, S2 and S3):

$$CC = \sum_{i=1}^{4}\left(\sum_{j=1}^{4} w_{ij} * P(C_j/E_i) * P(E_i)\right) \tag{7}$$

Where:

$E_i \epsilon \{E_1, E_2, E_3, E_4\}$ is the indicator grade according to experts;

$C_j \epsilon \{C_1, C_2, C_3, C_4\}$ is the indicator grade based on a comparison (equation 6) between the density score *D* and a tested set of thresholds (S1, S2, S3);

$P(C_j/E_i)$ are the probability of assigning to grade C_j a segment assessed by expert into grade E_i ;

$P(E_i)$ are the probability that a segment is in grade E_i. This probability depends on the study population.

w_{ij} are the weights attributed to an error: segment assigned wrongly to grade j according to the threshold values but to grade i according to the expert's opinion (as reference).

- If $i = j$ then $w_{ij} = 0$.
- If i<j the error is a false positive and dysfunction is overestimated by comparing score and thresholds.
- If j<i the error is a false negative and dysfunction is underestimated by comparing score and thresholds.

w_{ij} depends on the capacity of the utility manager to manage existing risks related to dysfunctions. w_{ij} is a 4x4 matrix as shown by table 5-4. A false negative error can have major consequences due to underestimating the real condition of the sewer. It is therefore rational to assign higher values to all w_{ij} when ($i>j$). On the other hand, the consequence of a false positive may be an unnecessary investigation of the segment and smaller values should therefore be used for all w_{ij} when ($i<j$).

Table 5-4. Calibration weights matrix (w_{ij})

Expert evaluation	**Grade calculated from single score**			
	G1	**G2**	**G3**	**G4**
G1	0	w_{12}	w_{13}	w_{14}
G2	w_{21}	0	w_{23}	w_{24}
G3	w_{31}	w_{32}	0	w_{34}
G4	w_{41}	w_{42}	w_{43}	0

This criterion makes use of three types of information. $P(C_j/E_i)$ are determined from a set of sewer segments, known as the "national sample", assessed by a national working group (engineers in charge of asset management in several French utilities). Section 5.4.3 describes the national sample.

$P(E_i)$ and w_{ij} are related to the local asset. If visual inspections are carried out for the purpose of a sample survey (code ABP J under the EN 13508-2 coding system) then this representative sample can be used to assess $P(E_i)$. This case is presented in section 5.4.5. If not, $P(E_i)$ are taken as hypotheses (section 5.4.3).

5.4.3. National sample

For each dysfunction indicator (table 5-1), expert assessments are collected and used in order to determine the optimum set of thresholds. Experts evaluated the national sample, made up of a set of 60 segments within three French utility areas. More than 50% of these segments belong to the asset stock of the Greater Lyon utility and the other segments are from Strasbourg and Bas-Rhin. These segments are evaluated from their CCTV reports under EN13508-2 including existing defect pictures.

Expert opinions are gathered in such a way that each dysfunction indicator for each segment is evaluated at least by four different experts. This means that for each dysfunction indicator from each segment, there are four or more answers.

In some cases, an expert is undecided between grade *i* and *i-1*, therefore a hybrid grade for this segment (grade *i/i-1*) is considered. In some cases, all the experts do not agree on a specific grade. Werey *et al.,* (2008) and Cherqui *et al.,* (2008) propose a method to deal with this problem. By using this method for correcting experts' opinions, the distribution of expert assessments versus calculated scores for the infiltration indicator (INF) is presented in Figure 5-3. The assigned values of *a* and *b* (defined in section 5.3.2) are respectively 3 and 1. Obviously, this figure is not an ideal distribution of expert assessments versus scores (compared to figure 5-2). Some differences of interpretation between experts are due to their backgrounds, as discussed in Werey *et al.,* (2008). Dirksen *et al.,* (2011) show that interpretation of a CCTV report may introduce some uncertainties.

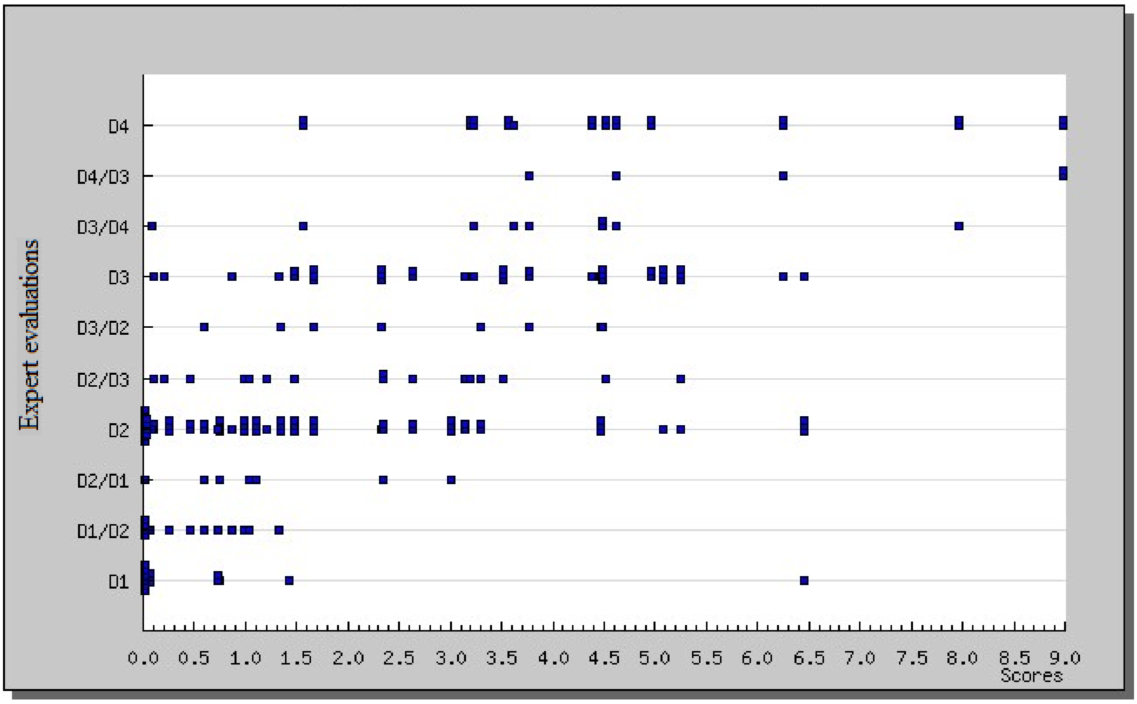


Figure 5-3. Distribution of expert evaluations vs. calculated scores (a=3 and b=1)

5.4.4. Calibration without a representative sample of the asset stock

This procedure is used if the inspected segments do not constitute a representative sample of the studied asset stock. This means that inspections are carried out in particular areas and/or according to specific goals (for example code ABP F: suspected infiltration problem or ABP D: suspected structural problem under the EN 13508-2 coding system). In this situation, the calibration process based on criterion introduced in section 5.4.2 uses two sources of information:

- The national sample.
- Hypotheses about the global condition of the studied asset stock.

As mentioned above, $P(E_i)$ represents the global condition of the asset (the proportion of segments in each grade). In the case where no representative sample is available, these proportions are put forward by the utility managers, depending on their knowledge about the general condition of the asset stock. For example, an asset stock in a good general condition may correspond to $P(E_1)$ = 30%, $P(E_2)$ = 50%, $P(E_3)$ = 15% and $P(E_4)$ = 5% and an asset in a bad general condition may correspond to $P(E_1)$ = 15%, $P(E_2)$ = 15%, $P(E_3)$ = 35% and $P(E_4)$ = 35%.

On the other hand, the calibration criterion *CC* also depends on assignment-error weights w_{ij}. Thresholds consequently depend on these weights that relate to false positive and false negative errors. These weights represent the intentions chosen by wastewater services and represent also their practices (actions when a segment is classified in grade C_j).

The calibration process uses a systematic procedure exploring a 3D solution space. A possible solution verifies that $0 \leq S_1 < S_2 < S_3 \leq D_{max}$ (with $D_{\max}$ the maximum value of the single score within the national sample). The calibration criterion *CC* is calculated for each possible triplet of thresholds (*S1*, *S2*, *S3*) (equation 7). For a *k* set of thresholds $(S1, S2, S3)_k$ applied to the national sample, conditional probability that a segment in grade E_i is assigned to grade C_j is computed as follows:

$$P(C_j/E_i) = \frac{NB(C_j/E_i)}{NB(E_i)} \tag{8}$$

Where $NB(Cj/E_i)$ represents the number of segments assigned to grade *j* according to the threshold values and to grade *i* according to an expert's opinion, $NB(E_i)$ represents the total number of segments assigned to grade *i* according to an expert's opinion. Hence, the required values of $P(C_j/E_i)$ depend only on thresholds. The triplet which minimizes the calibration criterion is then assigned as threshold values.

In section 5.5, this method is applied to our national sample by considering different values for w_{ij} and $P(E_i)$. The calculated thresholds will then be tested on CCTV reports of the Greater Lyon utility. The effects of using these different values will be discussed at the end of this chapter.

5.4.5. Calibration with a representative sample of the asset stock

If a representative sample of an asset stock is carried out from specific inspection campaigns within a utility, the calibration should therefore be more specific to the utility's asset. The calibration process is an iterative procedure and the calibration criterion *CC* will be calculated for all the different sets of thresholds (*S1*, *S2*, *S3)*. However, $P(E_i)$ now represents the condition of the studied population and thus its determination is based on the distribution of sewer segments (in the representative sample) in each grade (depending on a *k* set of thresholds (*S1*, *S2* and $S3)_k$) (Figure 5-4).

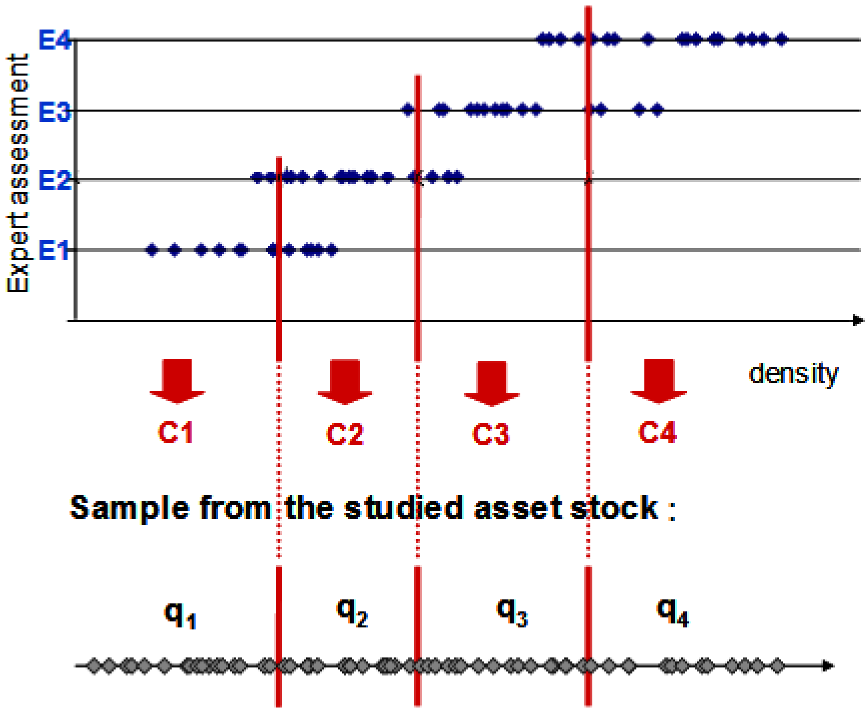


Figure 5-4. Calibration procedure with a representative sample of the asset

Figure 5-4 shows the use of a set of thresholds (vertical lines) on a sample of assets from the studied asset stock. Assuming that $NB(C_j)$ is the number of segments of the representative sample assigned to grade C_j then q_j represents the proportion of segments assigned to a grade C_j as follows:

$$q_j = \frac{NB(C_j)}{\sum_{j=1}^{4} NB(C_j)} \tag{9}$$

We denote α_{ij} the proportion of segments in grade E_i which is assigned in grade C_j. As for the previous process, the national sample provides α_{ij} for a given set of thresholds:

$$\alpha_{ij} = \frac{NB(C_j/E_i)}{NB(E_i)} \tag{10}$$

q_j (j=1 to 4) and α_{ij} are successively the arrays of column matrix of $\boldsymbol{q}$ and square matrix of $\boldsymbol{A}$. We also define $\boldsymbol{P}$ as a column matrix containing all $P(E_i)$.

Concerning the sample of the studied population, among the segments in grade E_i (that would have been assigned to grade E_i by experts), some of the segments are assigned to grade C_1, C_2, C_3 and C_4. Thus, relations between q_j and $P(E_i)$ can be written as follows:

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} & \alpha_{41} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \alpha_{42} \\ \alpha_{13} & \alpha_{23} & \alpha_{33} & \alpha_{43} \\ \alpha_{14} & \alpha_{24} & \alpha_{34} & \alpha_{44} \end{bmatrix} \begin{bmatrix} P(E_1) \\ P(E_2) \\ P(E_3) \\ P(E_4) \end{bmatrix} \tag{11}$$

So:

$$\boldsymbol{q} = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{P} \tag{12}$$

And at the end:

$$\boldsymbol{P} = (\boldsymbol{A}^{\boldsymbol{T}})^{-\boldsymbol{1}}\boldsymbol{q} \tag{13}$$

The values of w_{ij} are also to be set by utility managers.

This procedure is iterative: it begins with a hypothesis about vector P (P_0) (preferably $P(E_1)=P(E_2)=P(E_3)=P(E_4)=25\%$) and an initial set of thresholds $(S1,S2,S3)_{initial}$ verifying $0 \leq S_1 < S_2 < S_3 \leq D_{max}$ (for example (0.25, 1, 4)) used to calculate the first triplet of thresholds by minimizing the calibration criterion (equation 7). This new set of thresholds is then used to calculate ***q*** and ***A*** as both of them depend on threshold values. Then, new vector ***P*** is obtained using equation 13. This procedure iterates until the convergence of ***P***.

5.5. Sensitivity analysis of calibration parameters and application to the Greater Lyon sewer asset stock

A full-scale application of the method was realized in close collaboration with the Greater Lyon utility. This application seeks to assess the condition grade of all inspected sewer segments under EN13508-2. The inspected segments do not constitute a representative sample of the studied asset stock. Hence, we use the calibration procedure based only on the national sample explained in section 5.4.4.

A sensitivity analysis (section 5.5.1) is carried out with three adjacent hypotheses of overall condition of asset and three contrasted assignment-error weight matrices. The nine obtained scales of numerical values are used to assess the condition grade of the Greater Lyon sewer segments. Section 5.5.2 describes the studied asset stock. Section 5.5.3 summarizes the results of these evaluations.

5.5.1. Sensitivity analysis and calibration of thresholds

In section 5.4.2, a calibration procedure is proposed in order to obtain the scale of numerical values when a representative sample of the studied asset stock is not available. However, two elements of equation 7 need further investigation:

(1) $P(E_i)$ is the probability that a segment is in grade E_i. When a representative sample of the asset stock is not available, the values of $P(E_i)$ are estimated by the utility manager. However, the calibration process is affected by these values. For example, an increase in $(P(E_3)/P(E_4))$ will lead to a decrease of FP numbers (segments in G3 according to expert evaluations and in G4 according to threshold values). The value of S3 (threshold between G3 and G4) will therefore be increased. Three hypotheses

regarding the condition of the studied asset stock are proposed for sensitivity analysis as no representative sample of the Greater Lyon asset stock exists.

(2) wij (weight given to an assignment error) also influences the threshold values. For example, an increase in (w_{43} / w_{34}) gives greater weight to false negative (FN) errors rather than false positive (FP) errors. This means that the number of FN errors decreases. Three matrices of weights are also compared.

The different hypotheses about $P(E_i)$ and w_{ij} are described in table 5-5 and figure 5-5.

Table 5-5. Three hypotheses (*P1*, *P2* and *P3*) about the probability $P(E_i)$ that a segment is in grade G_i

Hypothesis on the asset stock	Condition grades			
	G1	G2	G3	G4
P1	10%	40%	40%	10%
P2	30%	30%	20%	20%
P3	20%	30%	30%	20%

M1

0	1	3	5
1	0	1	3
3	1	0	1
5	3	1	0

M2

0	1	1	1
1	0	1	1
5	2	0	1
10	5	3	0

M3

0	1	1	1
3	0	1	1
15	6	0	1
30	15	9	0

Figure 5-5. Three matrices of assignment-error weight

Three hypotheses about $P(E_i)$ combined with three matrices of w_{ij} lead to nine evaluations (scenarios) (table 5-6).

Table 5-6. Definition of scenarios

Scenario	J1	J2	J3	J4	J5	J6	J7	J8	J9
Matrix/distribution	M1/P1	M1/P2	M1/P3	M2/P1	M2/P2	M2/P3	M3/P1	M3/P2	M3/P3

P1, *P2* and *P3* correspond to three distributions of G1, G2, G3 and G4. These sets represent an asset in bad condition (50-60% of the network in G1+G2 and 40-50% in G3+G4). Sensitivity analysis results presented in figure 5-6 show that for a given assignment-error weights matrix, threshold values do not vary significantly between P2 and P3. Hence, they do not vary significantly when distributions represent a similar condition of the studied asset stock.

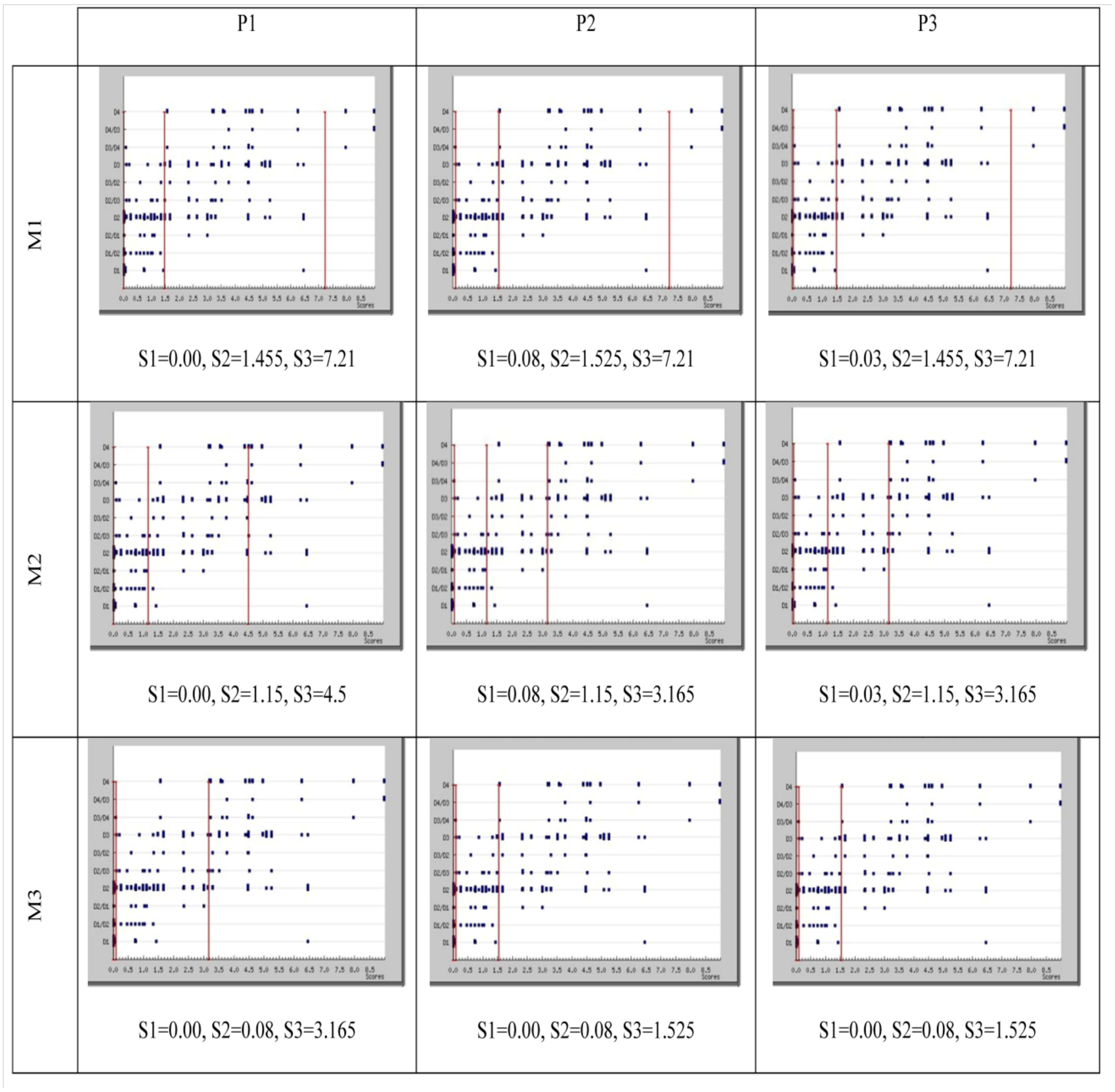


Figure 5-6. Results of sensitivity analysis: 3 hypotheses on the asset stock (P1, P2, P3) combined with 3 hypotheses on assignment-error weights (M1, M2, M3).

M1 is a symmetric matrix. An equal value for w_{ij} and w_{ji} is attributed in order to assign the same weight to an error, regardless of whether it is a FP or FN error. Using this matrix means that overestimating or underestimating errors are the same for the asset manager. In M2 and M3 matrices, a constant value equal to one is attributed to all the FP errors. The FN values increase as the difference between the calculated grade according to threshold values and from expert evaluation aggravates.

The influence of using higher values for FN cases is studied by tripling all values of FN cases in M3 comparing with M2 which makes the former a more-secure matrix. Consequently, we expect a higher proportion of segments in G3 and G4 for J7, J8 and J9.

Results presented in figure 5-6 show that by increasing the weights of FN errors, threshold values decrease (for any $P(E_i)$). Hence, threshold values are sensitive to the weights given by the utility managers.

We can therefore conclude that:

(1) By increasing the weights of FN errors, threshold values decrease. This is true for any P(Ei).

(2) By considering small weights for all FN and FP errors, the assumed condition of the studied asset stock P(Ei) does not have a significant influence on threshold values (M1 case).

(3) The P(Ei) gain more importance by increasing the weights of FN errors. For example, threshold values vary significantly between P1 and P2 or P3 for M2 and M3.

In section 5.5.3, we present the consequences of using each of these scenarios on a set of segments from Greater Lyon. Finally, it should be mentioned that the influence of the national sample on the results is unidentified. Therefore further investigations will be carried out regarding the influence of national sample size following new surveys.

5.5.2. Study sample from the Greater Lyon asset stock

A set of 4,471 segments (152km) is available for comparison with the results of the aforementioned scenarios. These segments, inspected according to EN 13508-2, belong to the asset stock of the Greater Lyon urban community. Their lengths vary from 5-70m. Figure 5-7 shows the cumulative percentage of segments versus their single scores.

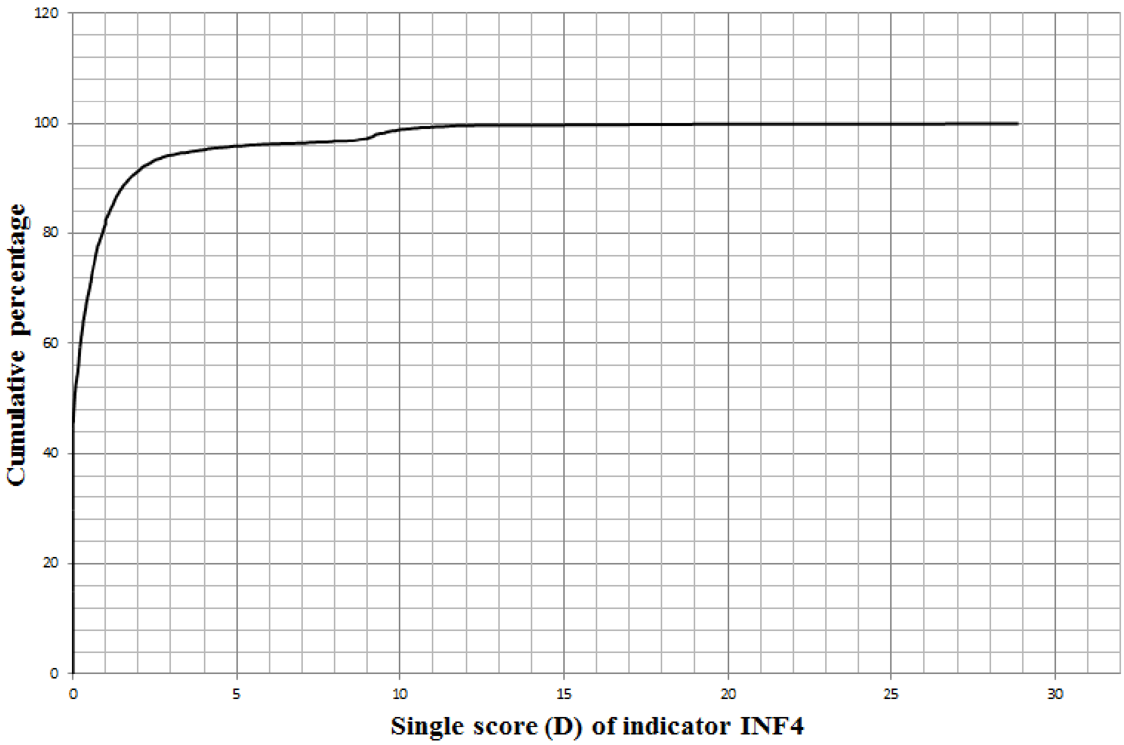


Figure 5-7. Cumulative percentage of segments versus the single score

5.5.3. Result

The results of applying the scenarios mentioned in table 5-6 for the infiltration indicator (INF) are presented in table 5-7 (number of segments) and table 5-8 (as a percentage). As we expected, J1 to J6 show more or less same results, as about 85% and 15% of segments are respectively in the G1-G2 and G3-G4 condition grades. In other words, the results do not vary significantly from one scenario to another, although almost all the thresholds are changed. On the other hand, the variation in the number of segments in G3-G4, being the critical part of network, is only 4% between these scenarios.

More segments are in the critical condition grades (G3 & G4), with the J7, J8 and J9 scenarios. These scenarios use M3 weight matrix which is more secure than M1 or M2. Hence, 48% of all segments are assigned to the G3-G4 group of condition grades. This occurs as a result of using high values for FN assignment error weights. As expected, the assignment-error weight matrix plays a very important role in the calibration process and has a significant impact on the results.

Table 5-7: Number of segments in different condition grades

	Scenario								
Grade	J1	J2	J3	J4	J5	J6	J7	J8	J9
G1	2040	2320	2138	2040	2320	2138	2040	2040	2040
G2	1892	1640	1794	1731	1451	1633	280	280	280
G3	383	355	383	503	453	453	1904	1640	1640
G4	156	156	156	197	247	247	247	511	511
Total	4471	4471	4471	4471	4471	4471	4471	4471	4471

Table 5-8: Percentage of segments in different condition grades

Weights matrix	Condition grade	Hypotheses on the asset stock		
		P1	P2	P3
M1	G1	46%	52%	48%
	G2	42%	37%	40%
	G3	9%	8%	9%
	G4	3%	3%	3%
M2	G1	46%	52%	48%
	G2	39%	33%	37%
	G3	11%	10%	10%
	G4	4%	5%	5%
M3	G1	46%	46%	46%
	G2	6%	6%	6%
	G3	43%	37%	37%
	G4	5%	11%	11%

On the other hand, a more in-depth comparison between J2 and J5 and also J4 and J5 leads us to a better understanding of the results obtained (table 5-9 & 10).

Table 5-9. Comparison between J2 and J5

Result with J2	Result with J5			
	G1	G2	G3	G4
G1	2320 (52%)	0	0	0
G2	0	1451 (32%)	189 (4%)	0
G3	0	0	264 (6%)	91 (2%)
G4	0	0	0	156 (3%)

For example, 280 segments which are initially assigned into G2 with scenario J4, are assigned into G1 with scenario J5, with this being due to an increase of S1 from 0 (J4) to 0.08 (J5). In addition, as S3 decreases from 4.5 (J4) to 3.165 (J5), 50 segments assigned into G3 with J4 are allocated into G4 with J5. These results show the influence of $P(E_i)$ on the assigned condition grade.

Table 5-10. Comparison between J4 and J5

Result with J4	Result with J5			
	G1	G2	G3	G4
G1	2040 (46%)	0	0	0
G2	280 (6%)	1451 (32%)	0	0
G3	0	0	453 (10%)	50 (1%)
G4	0	0	0	197 (4%)

Another example is the comparison between J2 and J5 in order to study the influence of assignment-error weights. As S2 decreases from 1.525 (J2) to 1.15 (J5), 189 segments, initially affected into G2 in J2, are shifted into G3 in J5. It is important to note that:

(1) More than 90% of segments stay in the same condition grades (diagonal arrays) with the J2-J5 and J4-J5 scenarios presented in table 5-6;

(2) By comparing these scenarios, a segment's condition grade varies only between 'adjacent' grades (G1 with G2, G2 with G3 and G3 with G4);

(3) The changes in the non-critical part of the asset (G1-G2) are not very important. In other words, there are no significant consequences if a segment is assigned to G1 by a scenario and to G2 by another scenario (J4-J5 case). On the other hand, consequences can be very significant if this happens between G2 and G3 (J2-J5 case) and G3 and G4 (both cases);

(4) Only 1% of all segments are assigned into G3 with J4 and into G4 with J5, but this proportion is 20% of segments in G4 with J5. It is true that 40% of segments

allocated to G4 with J5 are assigned into G3 with J2, but this accounts for just 2% of all segments.

The synthesis of such sensitivity analyses can be used to prioritize investigation and/or rehabilitation work. For instance, segments which are systematically assigned into G4 by all scenarios are the most critical. Segments whose condition grade changes from G4 into G3 constitute a secondary set of candidates.

5.6. Conclusion

An important phase of sewer asset management is the evaluation of physical condition of assets and the assessment of a condition grade to a sewer segment. While condition assessment of sewer segments relies mostly on visual inspections, each condition grading protocol uses a different approach for taking into account observed defects.

The RERAU methodology uses the notion of mean score. A strong linear relationship between total score and segment length confirms the ability of mean score to characterize a segment for a particular dysfunction indicator. We showed that a strong linear relationship exists between the total score of infiltration indicator and segment length.

While most protocols assign an indicator into a grade by comparing this mean score with a subjective scale of numerical values, this paper proposes a method to determine thresholds for each specific asset stock. This method has 3 specific advantages:

- A calibration criterion is optimized in order to justify the scale of numerical values;
- This criterion takes into account local specificities (as criterion parameters);
- A sensitivity analysis allows tackling the uncertainty linked to the parameters.

The proposed calibration criterion allows threshold calculation according to two sets of parameters: overall condition of the asset stock in question and assignment-error weights associated to false positive and false negative errors. The former parameter is either taken as a hypothesis or is determined by a representative sample of studied asset stock while the latter parameter is fixed by the utility manager. This method relies also on a national sample, including expert assessments of a set of selected sewer segments.

This protocol led to the development of a software tool. The application of this software on 4,471 segments of the Greater Lyon sewer network is reported in this chapter (for a total length of 152km).

This proofs that local specificities must be taken into account because they influence the scale of numerical values used to assign an indicator into a grade.

The results of sensitivity analysis of model parameters show that the calibration process is sensitive to the parameters (condition of the asset stock and weights associated to assignment errors). These weights may be defined almost *ex nihilo* as they depend on the strategy of the utility manager. However, knowledge about the overall condition of the asset stock requires special inspection surveys in order to put together a representative sample.

There are two main areas of future research:

- Improvement of the calibration requirements:
  - Increasing the size of the national sample
  - Replacing the discrete assessments of experts by a continuous distribution for each grade (Leeflang *et al.,*2008)
- Defining method(s) to precisely assess the overall condition of an asset stock and more specifically a method to define a representative sample (addressed in chapter 6);

5.7. References

1. AGHTM (1999). Les ouvrages d'assainissement non visitables : fiches pathognomoniques. Techniques Sciences et Méthodes, octobre, 23-90 (In French).
2. Ahmadi M., Cherqui F., De Massiac J.C., Werey C., Lagoutte S., Le Gauffre P., (2012). "Condition grading for dysfunction indicators in sewer asset management". *Journal of Structure and Infrastructure Engineering*. (DOI: 10.1080/15732479.2012.756916).
3. ATV-M 143 Teil 2 (1999). Inspektion, Instandsetzung, Sanierung und Erneuerung von Abwasserkanälen und leitungen – Teil 2: Optische Inspektion. Hennef: GFA, April 1999.- ISBN 3-933693-43-8
4. Alegre, H., Hirner, W., Baptista, J. M., and Parena, R., 2000. "Performance indicators for water supply services." IWA (International Water Association) manual of best practice, IWA Publishing, London.
5. Ana, E. and Bauwens, W., 2010. "Modeling the structural deterioration of urban drainage pipes: the state-of-the-art in statistical methods. Urban Water Journal, 7(1), 47 – 59.
6. CEN, 2003, EN 13508-2 Condition of drain and sewer systems outside buildings, Part 2: Visual inspection coding system, European Norms.
7. Cherqui, F., Werey, C., Ibrahim, M., Le Gauffre, P., 2008. CCTV inspection of sewer segments: calibration of performance indicators based on experts' opinions. Proceedings of the 11th International Conference on Urban Drainage, Edinburgh, UK, 31 Aug.-5 Sept., 10 p.
8. Dirksen, J., Clemens, F. H. L. R., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K., Snaterse, C. T. M., 2011. The consistency of visual sewer inspection data, Structure and Infrastructure Engineering, First published on: 07 February 2011 (iFirst).
9. Ellis, J.B. and Bertrand-Krajewski, J.-L. (Ed.), 2006. "Assessing infiltration and exfiltration on the Performance of Urban Sewer Systems (APUSS)". IWA Publishing, London, 192 p.
10. Hahn, M., Palmer, R., Merrill, M., Andrew, B., Lukas, A., 2002. Expert system for prioritizing the inspection of sewers: Knowledge base formulation and evaluation. ASCE, Journal of Water Resources Planning and Management, 128(2), 121-129.
11. Ibrahim, M., Cherqui, F., Le Gauffre, P., Werey, C., 2007. Sewer asset management: from visual inspection survey to dysfunction indicators. LESAM 2007 – 2nd IWA Leading Edge Conference on Strategic Asset Management, International Water Association. Proceeding on CD-Rom. Lisbon 17-19 oct. 10 p.
12. Knolmar, M. and Szabo, G. C., 2003. CARE-S report D3 - WP2 Structural condition - Classification systems based on visual inspection, Technical report of the CARE-S project, Budapest University of Technology and Economics, 192 p.
13. Le Gauffre, P., Cherqui, F., 2009. Sewer rehabilitation criteria evaluated by fusion of fuzzy indicators. LESAM 2009 – 3rd IWA Leading Edge Conference on Strategic Asset Management. Miami, U.S.
14. Le Gauffre, P., Joannis, C., Vasconcelos, E., Breysse, D., Gibello, C., Desmulliez, J.J., 2007. Performance Indicators and Multicriteria Decision Support for Sewer Asset Management. Journal of Infrastructure Systems, ASCE, 13(2), 105-114 (in French).

15. Le Gauffre, P., Joannis, C., Breysse, D., Gibello, C., Desmulliez, J.J., 2004. Gestion patrimoniale des réseaux d'assainissement urbains. Guide méthodologique. Paris, Lavoisier Tec&Doc, 416 p., ISBN 27430-0748-6.
16. Leeflang M.M., Moons K.G., Reitsma J.B., Zwinderman A.H., 2008. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: Mechanisms, magnitude, and solutions. Clin Chem. 54:729-37. [PMID:18258670]
17. Marlow, D. and Burn, S., 2008. Effective use of Condition Assessment within Asset Management, JAWWA, 100(1), 54-63.
18. Marlow, D., Heart, S., Burn, S., Urquhart, A., Gould, S., Anderson, M., Cook, S., Ambrose, M., Madin, B., Fitzgerald A., 2007. "Condition assessment strategies and protocols for water and wastewater utility assets", WERF - Water Environment Research Foundation, IWA Publishing, 476 p.
19. Matos, R., Ashley, R.M., Cardoso, A., Molinari, A., Schulz, A., Duarte, P., 2003. Performance indicators for wastewater services. International Water Association, ISBN: 1900222906.
20. Merrill, S., Lukas, A., Palmer, R.N., Hahn, M.A., 2004. Development of a tool to prioritize sewer inspections. Technical report, WERF - Water Environment Research Foundation, 123 p.
21. Rahman, S. and Vanier, D.J., 2004. "MIIP Client Report: An Evaluation of Condition Assessment Protocols for Sewer Management", National Research Council of Canada, Ottawa.
22. Saegrov, S. (Ed.), 2006. "CARE-S - Computer Aided Rehabilitation for Sewer and Stormwater Networks", IWA Publishing, ISBN: 1843391155, 140 p.
23. Saegrov, S. (Ed.), 2005. "CARE-W: Computer Aided Rehabilitation for Water Networks", IWA Publishing, London, 208 p.
24. Vanier, D.J, Newton, L.A., Rahman, S., 2006. A Framework for municipal infrastructure management for Canadian municipalities. Building Science Insight conference.
25. Vanier, D.J. and Newton, L.A., 2006b. MIIP Report: The State of Canadian Sewers - Analysis of Asset Inventory and Condition. Client Report. NRC Institute for Research in Construction, 38 p.
26. Vanier, D.J. and Rahman, S., 2004. MIIP Report: A Primer on Municipal Infrastructure Asset Management. Client Report. NRC Institute for Research in Construction, 71 p.
27. Vanier, D.J., 2001. Asset Management: A to Z, NRCC/CPWA/IPWEA Seminar Series, Innovations in Urban Infrastructure, Philadelphia, PA, 1-16.
28. Water Research Center (WRc). 2004. Sewerage rehabilitation manual, 4th Ed., WRc, Wiltshire, U.K.
29. Werey, C., Cherqui, F., Ibrahim, M., Le Gauffre, P., 2008. Sewer asset management tool: dealing with experts' opinions. In Cabrera & Pardo (ed.) Performance Assessment of Urban Infrastructure Services. Drinking water, waste water and solid waste. London, UK: IWA Publishing, 125-134.
30. Werey, C., Dorchies, D., Mellac Beck, I., 2006. Sewer asset management: assessing criteria for a multicriteria decision support on a county level data base. Joint int. conf. on Computing and decision making in civil and building engineering, Montreal, June 14-16, 886-895.

Chapter VI: Assessment of an asset stock from a sample

This chapter is extracted from a paper entitled "Sewer asset management: impact of sample size and its characteristics on the calibration outcomes of a decision-making multivariate model" submitted in Water Research Journal (Ahmadi et al., submitted)).

6.1. Overview

According to the American Society of Civil Engineers, the current average condition grade of the national wastewater asset stock is *D* below the average of *D+* for all infrastructure systems in the United States using a simple grading scale from *A* to *F* (ASCE 2013). These evaluations have been being carried out since 1988. According to the former, the overall condition of assets aims at evaluating the infrastructure's in-time or near future physical condition. Along with the last evaluation, according to the prediction carried out by the U.S. Environmental Protection Agency (EPA) by 2020, 44% of sewer segments will be in a condition grade that necessitates an imminent action (Allbee, 2005). These evaluations are based on samples from the whole U.S. asset stock.

In fact, at the moment small number of utilities has completely inspected their asset stocks. Therefore, the use of a sample from an asset stock in order to calibrate deterioration models, to study scenarios about future and to draw appropriate conclusions seems mandatory. This sample should reflect, however, the characteristics of the asset stock in-question in the best manner. By definition, a sample which is an appropriate image of an asset stock is the representative sample of this asset stock (Cochran, 1977; Lohr, 2010).

In addition, in chapter 5, we proposed an evaluation methodology in order to assign segments into an ordinal grade of G1-G4 (G1 the best condition and G4 the worst condition) based on

CCTV inspections. This methodology requires two parameters: 1) stakeholders' intentions and 2) the overall condition grade of the asset stock. In this case, having a representative sample of the asset stock or having some reliable guess about the overall condition of an asset stock is mandatory.

Baur and Herz (2002) forecast condition of sewers from a supposed representative sample of inspected sewers of Dresden city, Germany by using transition functions based on cohort-survival functions. However, the use of a representative sample of an asset stock is not just limited to calibration of deterioration models or to provision of the overall condition of an asset stock. In addition, in chapter 4, by using the deviance statistic (a likelihood ratio test), we propose a method allowing the establishment of the list of most informative single factors in terms of inspection program efficiency (finding segments in failure state). This method can be applied on a representative sample of the asset stock (containing segments' influential factors such as age, material etc. and their condition grades) in order to give utilities important recommendations about data acquisition plans responding to following question: *"what data to gather considering its importance and cost of acquisition"*.

However, within the scientific literature dedicated to the asset management, authors developed and calibrated deterioration models without paying attention to the impact of used sample on the outcomes. The main interest of doing so was to test and compare the deterioration models (Trans 2007) or to develop a risk-based rehabilitation approach (Salman and Salem 2012).

Baur and Herz (2002) calibrated their cohort transition functions by using only 2.7% of the total length of Dresden network as a representative sample of this network. In addition, they used quota sampling method to carry out this representative sample. Quota sampling depends strongly on the user judgment and is not a probability sampling method (Cochran, 1977).

For instance no study tackles with the problem of representativeness of the used samples in order to calibrate deterioration models (or transition functions as well). Hence, attention should be paid to following issues:

(1) The fact that we use just a sample of the asset stock. In other words, we should consider the following question: "Are our conclusions drawn from a sample generalizable to the whole asset stock?". For example Davis *et al.*, (2001b) by calibrating the binary logistic regression on a sample of the asset stock of Thames Water, UK, conclude that age, depth, traffic type and road type are not significantly different from zero. They also comment that age is marginally insignificant and they remove it from the analysis which in next level, could be applied on the whole asset stock.

(2) What if we have not observed all patterns existing within the asset stock in our sample? For example, assume that the proportion of segments in brick within an asset stock is small and within our available sample any segment does not represent segments in brick. Hence, we cannot draw any conclusion for segments in brick from this specific sample in first place before passing into the next level which is generalization of findings to the whole asset stock. For instance no study tackles with the problem of representativeness of the used samples in order to calibrate deterioration models.

(3) However, depending on the deterioration model used and available sample, the calibration process may be problematic. Concato *et al.* (1993) quote that multivariable methods of analysis have been suspected of producing problematic results if too few outcome events are available relative to the number of independent variables analyzed in the model. For example in the case of using the logistic regression, three types of errors have been identified (Peduzzi *et al.*, 1996): over-

fitting (Type I error) occurs when too many variables (factors), some of which may be "noise," are selected for retention in the final model; under-fitting (Type II error) occurs when important variables are omitted from the final model; and paradoxical fitting (Type III error) is produced when a particular factor is given an incorrect direction of association which is the opposite of the true effect. Therefore, studies suggest a minimum number of events (failure state) per variable (EPV) about 10-20 EPV (Harrell *et al*., 1985; Peduzzi *et al*., 1995) in order to correctly calibrate the logistic regression coefficients. Bagley *et al*., (2001) compare the existing studies in this field and conclude that it is recommended that authors, reviewers, and editors pay greater attention to guidelines concerning the use and reporting of logistic regression models.

6.2. Scope of this chapter

Concerning the explained problems in the last section, the main scopes of this chapter could be summarized in the following questions:

- How to draw a representative sample of an asset stock which reflects in the best and appropriate manners the characteristics of this asset stock?
- How to provide a reliable estimation of a specific property of the asset stock from this sample such as the proportion of segments in failure state (condition grade 5 according to WRc for example)?
- Calibration of multivariate models seems somehow tricky. Therefore, what is the impact of used sample on the calibration outcomes of these multivariate models?

Lohr 2010 states that *"a good sample of a population will be representative in the sense that characteristics of interest in the population can be estimated from the sample with a known* *degree of accuracy"*. In order to address these questions, sample surveys should be carried out. According to Cochran (1977), in our case the main steps in a survey are as follows:

(1) Objectives of the survey: this step is directly linked to above questions as they do define the main purposes of the survey that is expected to be carried out.

(2) Population to be sampled: the word population is sometimes a controversial term in the designing phase of a survey. On the other hand, the sampled population should coincide with the population about which information is wanted (the target population). In our case, we make hypothesis that a list of segments is available within the utility database. Each segment has also a specific ID which will be used in the following steps.

(3) Data to be collected: normally during the surveys, all essential data should be gathered. In our case, during CCTV inspections of chosen segments' characteristics and environment should be gathered (size, gradient, depth, road class, etc.).

(4) Degree of precision desired: the results of sample surveys are always subject to some uncertainty as just one or some parts of population have been assessed.

(5) Methods of measurement: in the case where the objective of the survey is to provide an estimation of the proportion of segments in failure state, we normally use CCTV inspections. These evaluations are also subject to uncertainty (see chapter 5 and Dirksen et al., 2012).

(6) Sampling frame: the population must be divided into specific units. Depending on the used sampling method(s), each unit has a chance to be selected (probability sampling). This list of units is called the sampling frame. In our case, each segment constitutes a unit. According to Le Gauffre et al. (2007), a sewer segment corresponds to a length of several meters which is homogeneous in some

characteristics (material, diameter, etc.). In general, it is also delimited by two successive manholes.

(7) Selection of the sample: or sampling method is a variety of plans by which a sample may be selected. For each sampling method, estimation of the size of sample can be made from knowledge of the degree of precision desired.

(8) Organization of inspections: the mobilization of CCTV operators, our requested data and administrative aspects should be considered.

(9) Data analysis: depending on the objective(s) of survey, data gathered in last step should be analyzed and appropriate conclusions should be drawn.

The remainder of this chapter is organized as follows. In the next section, first the simulations' framework is explained. Statistics are then proposed in order to gauge the ability of defined simulations to estimate the characteristics of the population. Moreover in section 6.3.2 the sampling methods used in this chapter are discussed not only to estimate the minimum size of a randomly-chosen sample in order to be representative of the asset stock, but also to explain by which methods samples are drawn for simulations.

In section 6.4, we will describe our asset stock as the population. Finally, section 6.5 is dedicated to explanation of outcomes.

6.3. Methods

As explained in section 6.2, the main objective is to study the influence of sample used for the calibration of a multivariate deterioration model. In chapter 4, we showed the use of binary logistic regression in order to elaborate sewer inspection programs.

Assuming that K influential factors exist within a given database, the general form of the binary logistic regression is as follows:

$$L(x) = logit(Y = 1) = \ln\left(\frac{P(Y=1|X_1,X_2,X_3,\ldots,X_K)}{1-P(Y=1|X_1,X_2,X_3,\ldots,X_K)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_K \quad (1)$$

Where β_0= intercept term, $\beta_1, \beta_2, \ldots, \beta_K$= variable coefficients, $X_1, X_2, \ldots, X_K$= variables (influential factors) and Y = binary dependent variable (condition grade).

The probability of being in failure state is then obtained by:

$$P(Y = 1|X_1, X_2, X_3, \ldots, X_K) = \frac{1}{1+e^{-L(x)}} \quad (2)$$

The second objective is to draw a representative sample of the network and to study the behavior of the used multivariate model (binary logistic regression) when a representative sample is available. In section 6.3.2, we describe three frequently-used sampling methods in order to calculate the minimum size of a representative sample of the population (sewer network) derived randomly.

6.3.1. Simulations framework

Assume that *n* is the size of the sample used in order to calibrate the logistic regression (number of segments in the calibration sample). By applying the Monte Carlo method, for each sampling methods described in section 6.3.2, 1000 different samples of size *n* are drawn. Afterwards, for each sample, the logistic regression is fitted by using the maximum likelihood estimation method (Agresti, 2002) and the resulting coefficients and their standard errors are saved for an analysis described later in section 6.3.3. The simulations were performed by MATLAB®. A criterion of 10^{-5} was used for convergence of the maximum likelihood estimate. The sample size *n* is then varied from 600 to 5000 segments by a step of 400 segments.

This retrospective approach (Peduzzi *et al.*, 1995 & 1996) is used to vary the sample size *n* while at the same time leaving the regression coefficients apart from the intercept term unchanged (for the proof see Peduzzi *et al.*, 1995). However, theoretically if the used sample for the calibration is carried out randomly, once the sample size *n* is greater than the minimum required size for a sample to be representative (Cochran, 1997), the intercept term also remains unchanged. The retrospective approach gives the distribution of estimates of the logistic regression coefficients in relation to the sample size *n*.

6.3.2. Sampling methods: minimum size of a randomly-drawn sample in order to be representative of the population

In order to draw a representative sample of a sewer network, for simplicity reasons, we make following hypotheses:

- The interested variable is the proportion of segments in failure state. Failure state could be defined as the condition grade of segments which necessitate an imminent action (cf. chapter 5; Salman and Salem, 2012). In other words, our goal is to provide an accurate estimation of the proportion of segments in failure state.
- We suppose that every unit (segment) in the population (within the network), falls into one of the two classes F (in failure state) and F' (not in failure state).
- Material is known entirely for the population (all segments). We will use this information in order to stratify our population in section 6.5. We chose material because according to WERF (2009), each type of material has its specific deterioration pattern.

In the following section, some most important sampling methods regarding our objective are explained according to (Cochran, 1977 and Lohr, 2010). These methods are simple random sampling and stratified sampling by adopting proportional and optimum allocations. The minimum size of a randomly-drawn sample in order to be representative of the population is also given in each paragraph.

6.3.2.1. Simple random sampling (SRS)

Simple random sampling (SRS) is the most basic form of the probability sampling. A simple random sample without replacement of size n is selected so that every possible subset of n distinct units in the population has the same probability of being selected as the sample (Lohr 2010).

Assume that A is the number of units in class F in population and a is the number of units in class F in sample. Therefore, the proportion of units in class F in population is $P=A/N$ and in sample is $p=a/n$ where N and n are respectively the size of population and sample. Hence, the sample estimate of P is p and the sample estimate of A is Np or Na/n.

These estimations contain an amount of error depending on the characteristics of population. The tolerable amount of error depends on the purpose of survey and utility manager's intentions. In other words, considering an asset stock for which we guess or know from previous surveys that the proportion of segments in failure state is about 7%, fixing the estimation precision to 5% is not an appropriate decision. Similarly, for an asset stock with a guessed proportion of segments in failure state about 30%, fixing the estimation precision to 1% will be so costly.

Cochran (1977) proofs that the minimum sample size for obtaining a representative sample by assuming simple random sampling is provided by following equations:

$$n = \frac{n_0}{1+(\frac{n_0}{N})} \qquad (3)$$

Where $n_0 = \frac{pq}{V(p)}$ (4)

p is the proportion of units in F in sample, q is equal to (1-p), N is the size of population and V is the desired variance of the sample proportion.

According to Lohr (2010), if we expect that by (1-α)% of chance (α is the desired precision, see Cochran, 1977):

$P - e \leq p \leq P + e$, then:

$$V(p) = (\frac{e}{1.96})^2 \quad (5)$$

α is the desired precision and e is the margin of error. p (and consequently q) is estimated from previous surveys or is guessed by user.

6.3.2.2. Stratified random sampling

Stratification is a manner of giving more precision to our analysis if some conditions are met (Scheaffer *et al.*, 2012). Assume that population of N units is divided into L subpopulations of $N_1, N_2, \ldots, N_L$ units respectively ($N = \sum_{h=1}^{L} N_h$). Each subpopulation is called a stratum (in this chapter, each stratum corresponds to a different material). If a simple random sample is taken from each stratum (whose sizes are $n_1, n_2, \ldots, n_L$: $n = \sum_{h=1}^{L} n_h$), the whole procedure is described as stratified random sampling. In this chapter, we consider two stratified random sampling methods: proportional and optimum allocations (Cochran 1977).

One of our objectives is to estimate the proportion of the population that falls into class F for a given variable. We define:

$$P_h = \frac{A_h}{N_h}, p_h = \frac{a_h}{n_h} \quad (6)$$

These are respectively the proportions of units in class F in the population and in the sample for the stratum h (p_h is the sample estimate of P_h; A_h and a_h are the number of units in class F in the population and in the sample for the stratum h).

6.3.2.2.1. Proportional allocation in stratified random sampling (PSRS)

Proportional allocation uses a sampling fraction in each of the strata that is proportional to that of the total population. This means that for stratum *h*:

$$\frac{n_h}{n} = \frac{N_h}{N} = W_h \tag{7}$$

Where W_h is the *h*th stratum's weight. Similar to simple random sampling, for the minimum size of the sample *n* in order to be representative, Lohr (2010) shows that:

$$n_0 = \frac{\sum_{h=1}^{L} W_h p_h q_h}{V}, n = \frac{n_0}{1+\frac{n_0}{N}} \tag{8}$$

Where, similar to simple random sampling, *V* is the desired variance in the estimate of the proportion of *P* and *q=1-p*. Therefore, once *n* is assessed, n_h could be calculated by using equation 7. Moreover *V* is assessed by using equation 5.

6.3.2.2.2. Optimum allocation in stratified random sampling (OSRS)

In optimum allocation, each stratum is proportionate to the standard deviation of the distribution of the variable. According to Cochran (1977), the size of a given sample stratum increases if:

- The target population stratum has a large size,
- The stratum is more variable internally than the others;
- Sampling is cheaper for this stratum than other strata,

For this allocation, Cochran (1977) shows that the minimum sample size required to be representative could be calculated from the following equation if the sampling cost is identic within all strata:

$$n_0 = \frac{(\sum_{h=1}^{L} W_h \sqrt{p_h q_h})^2}{V}, n = \frac{n_0}{1+\frac{1}{NV}\sum_{h=1}^{L} W_h p_h q_h} \tag{9}$$

Similarly to simple random sampling and proportional allocation, V could be assessed by using equation 5. Hence, the size of each stratum is obtained from the following equation:

$$n_h = n\frac{N_h\sqrt{p_h q_h}}{\sum_{h=1}^{L} N_h\sqrt{p_h q_h}} \qquad (10)$$

6.3.3. Statistical analysis

As it is mentioned above, first, the regression model was fitted to the full sample (section 3). The values of coefficients calculated for the full sample (available asset stock) are called hereafter the *true* values. Afterwards, for each sample of size n, varying from 600 to 5000 segments by a step of 400 segments, 1000 Monte Carlo tries are carried out. This procedure is repeated for all sampling methods considered. By doing so, the distributions of the logistic regression coefficients in relation to the sample size n are obtained.

The simulation results of different sample sizes n were evaluated relative to the model fitted to the full sample or population. The accuracy, precision and statistical significance of each coefficient were also evaluated.

All statistics are computed under the condition of obtaining convergence of the logistic regression model. Samples in which the model did not converge were excluded from analysis. Although these samples provide some information about the parameter estimates (i.e., upper or lower bounds) they do not provide useful estimates of the effect of covariates (Peduzzi *et al*., 1996). These statistics are as follows:

(1) Accuracy of coefficients was studied by assessing the average relative bias for each of the regression coefficients ($k=1,\dots, K$) and each of the r simulations ($r=1, 2\dots R$) converged by using:

$$\sum_{r=1}^{R}(\beta_{kr} - \beta_{k,true})/\,(R\beta_{k,true}) \qquad (11)$$

$\beta_{k,true}$ is the *true* value of the coefficient k obtained from the population (full model).

(2) The precision of the estimated regression coefficients was expressed in standard errors (SEs). The SE was calculated for each coefficient of each simulation converged. Afterwards, the ratio of the average SE of coefficient k (over R converged simulations) and the *true* SE obtained from the population (full model) was assessed. This ratio approaches 1 when the SE is estimated correctly. This ratio is calculated by following equation:

$$(\sum_{r=1}^{R} SE_{kr}/\,R)/SE_{k,true} \tag{12}$$

(3) The statistical significance of the regression coefficients is evaluated in two ways. First, the proportion of simulations in which the 90% confidence interval does not meet zero and contains the *true* value. Second, the proportion of simulations in which the p-value for coefficient k is significantly different from zero (evaluating type I & III of the possible errors).

6.4. Databases

Our asset stock contains the data about age, material, sewer type, gradient, road class, depth, size, length and condition grade for each segment (table 6-1). The database contains 9810 segments. The total length of sewers is about 213 km. The factor "age" includes the effects of construction period and ageing. The average age of the asset stock is 80 years old. Approximately 70% and 23% of segments are made of vitrified clay and concrete respectively. About 22% and 25% of segments have a diameter equal to 200 and 300 millimeters respectively. The majority of segments are buried within a depth of 2 meters. The gradient for the majority of segments is less than 5%. Furthermore, about 70% of segments are buried under a street or an alley. Figure 6-1 shows the distributions of influential factors and segments condition grade.

Table 6-1: data available in our database

Data	Description	Nature of variable
Age	difference between the installation year of a segment and the year of inspection	Scale
Material	1: Brick, 2: vitrified clay, 3: clay tile, 4: reinforced concrete or 5: concrete	Categorical
Size	Diameter of the segment in millimeter	Scale
Depth	Average depth of buried pipe from the ground level	Scale
Gradient	Vertical displacement of the segment per horizontal displacement in percentage	Scale
Length	Length of segment in meter	Scale
Sewer type	1: Combined or 2: sanitary	Categorical
Road class	0, if the segment is not located under any type of roadway 1, if the segment is located under a street, an alley or a highway	Categorical
Condition grade	Binary structural condition grade of segment: 0- in good condition and 1- in failure state	Categorical

Our database is inspired from the asset stock of the metropolitan sewer district of Greater Cincinnati, USA (for more details see Salman 2010; chapter 4).

Table 6-2 provides the *true* values (the values obtained for the whole population) of the binary logistic regression coefficients $(\beta_0, \beta_1, \beta_2, \dots, \beta_K)$ using the maximum likelihood estimation method (Agresti, 2002). Vitrified clay, sanitary and road class=1 values are fixed as reference categories for material, sewer type and road class variables. Table 6-2 also shows the *p*-values and standard error (SE) for each variable. All of the regression coefficients for scale variables are significantly different from zero at a significance level of 0.05.

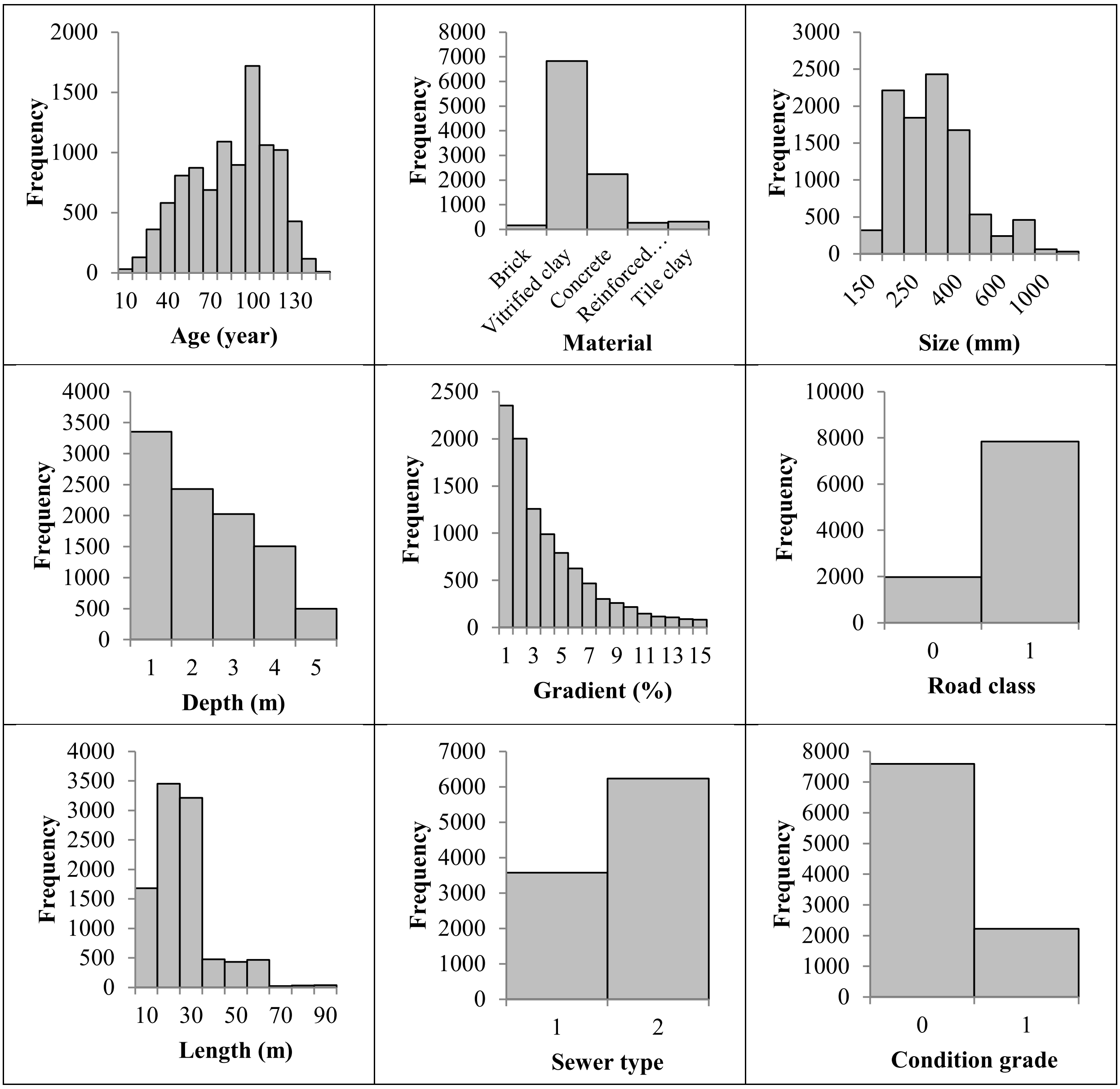


Figure 6-1: Histograms of certain influential factors and condition grade

Table 6-2: *True* values (binary logistic regression results) for asset stock 1

	Coefficient of the full sample or *true* values	*p*-Value	SE
intercept	-3.3531	0.0000	0.1356
Age (year)	0.0219	0.0000	0.0010
Material 1= Brick	-1.2229	0.0000	0.2822
Material 3= Concrete	-0.5569	0.0000	0.0814
Material 4= Reinforced Concrete	-0.4494	0.0224	0.1969
Material 5= Clay tile	-0.8370	0.0000	0.1751
Size (mm)	-0.0008	0.0002	0.0002
Depth (m)	-0.0422	0.0410	0.0206
Gradient (%)	0.0686	0.0000	0.0075
Road class=0	0.2678	0.0000	0.0608
Length (m)	0.0131	0.0000	0.0025
Sewer type= Combined	0.2457	0.0000	0.0534

Table 6-3 provides the corresponding parameters presented in section 6.3.2 in order to calculate the minimum size of a randomly-drawn sample to be representative of the asset stock.

Table 6-3: Sampling methods parameters of the population

Parameter	Description	Value
P	Proportion of segments in failure state in the population	22.6%
N	Number of segments in the population	9810
A	Number of segments in failure state in the population	2220
L	Number of strata, stratification is carried out according to materials	5
N_1, A_1	Total NO of segments and NO of segments in failure state in stratum 1	160, 16
N_2, A_2	Total NO of segments and NO of segments in failure state in stratum 2	6824, 1706
N_3, A_3	Total NO of segments and NO of segments in failure state in stratum 3	2245, 404
N_4, A_4	Total NO of segments and NO of segments in failure state in stratum 4	265, 53
N_5, A_5	Total NO of segments and NO of segments in failure state in stratum 5	316, 41

6.5. Results & discussions

6.5.1. Minimum required representative sample size

In order to draw a representative sample of the network, at the first place, we should define the precision (α) and the margin of error (e) by which we would like to calculate the minimum size of a randomly chosen segment. Assume we guess that the proportion of segments in failure state is around 20%. Therefore, we fix the margin of error to 3% and the desired precision to 0.05. In other words, we want, by 95% of chance, to have an estimate of P as follows:

$$P - 0.03 \leq p \leq P + 0.03 \tag{13}$$

p is the sample estimate of the proportion of segments in failure state (P).

For two allocations in stratified random sampling (proportional and optimum), we should make another hypothesis about the proportion of segments in failure state in each stratum. We consider the following scenarios (table 6-4):

(1) S1: Without any knowledge about ph. In this case we chose quantities which maximizes the n (the sample size).

(2) S2: An estimation of ph are provided by utility manager.

(3) S3: we know the real quantities of ph for all strata (not possible in practice before the survey).

Table 6-4: Different scenarios about p_h

Stratum	S1	S2	S3
1	50%	15%	10%
2	50%	30%	25%
3	50%	20%	18%
4	50%	25%	20%
5	50%	15%	13%

For simple random sampling:

$$V(p) = (\frac{e}{1.96})^2 = 2.34 \times 10^{-4} \rightarrow n = \frac{p(1-p)}{2.34\times10^{-4} + \frac{p(1-p)}{9810}} \qquad (14)$$

Figure 6-2 provides the sample size *n* as a function of an estimate of *P*. The maximum value of *n* is attained when *p* is equal to 0.5 ($n_{p=0.5} = 963$).

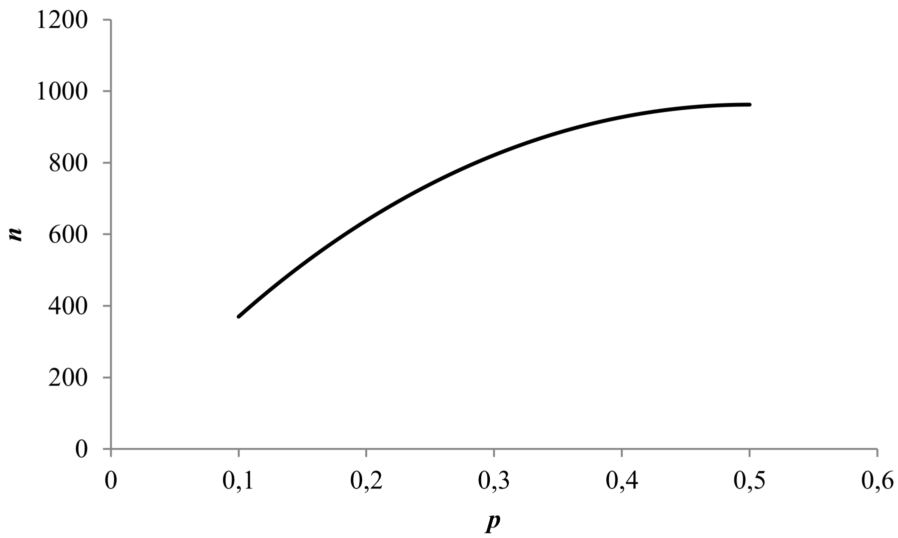


Figure 6-2: Size of the sample *n* in fonction of estimate proportion of segment in failure state *p* (*N*=9810)

For proportional and optimum allocation, table 6-5 and table 6-6 provides respectively the required total sample size and number of units (segments) from each stratum by considering defined scenarios above:

Table 6-5: Total sample size and number of segments from each stratum applying stratified random sampling

Stratum	Proportional allocation			Optimum allocation		
	S1	S2	S3	S1	S2	S3
1	16	12	11	16	10	8
2	670	531	480	670	551	497
3	220	175	158	220	158	145
4	26	21	19	26	20	18
5	31	25	22	31	20	18
Total	963	764	690	963	760	686

Following conclusions could be drawn:

- As expected the sample size for scenario 1 in both allocations in stratified random sampling is equal to the sample size obtained by simple random sampling for p=50%.
- By having a sample of 1000 randomly chosen segments, with all sampling methods, a representative sample may be drawn.
- As expected, in optimum allocation, the stratum with largest size should be more present within the sample (stratum 2).
- For other strata, the difference between proportional and optimum allocation in size is small.

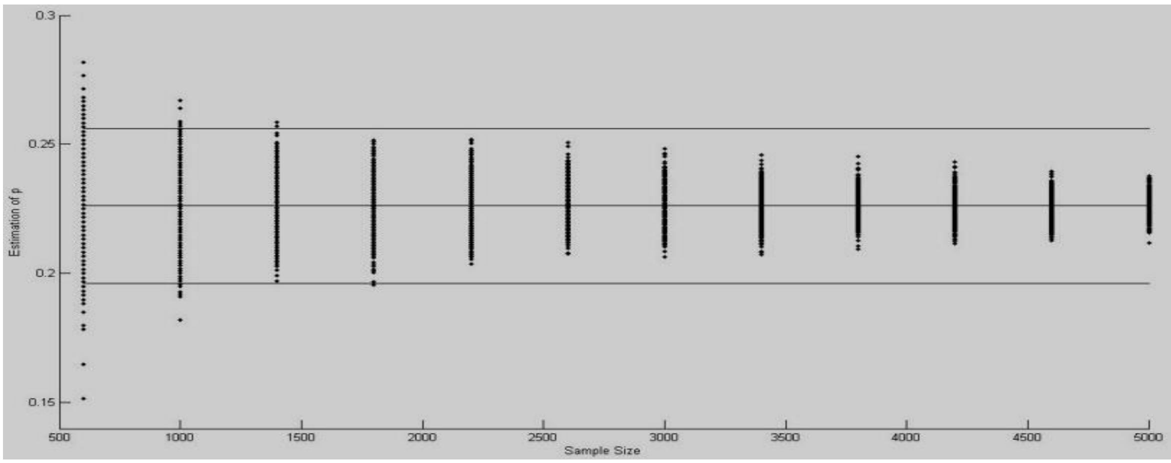
Figure 6-3 shows the results of 1000 simulations for each sampling method by varying n (the sample size) between 600 and 5000 segments with a step of 400 segments. For simple random sampling, these n segments are chosen completely randomly, though for proportional and optimum allocations in stratified random sampling, a specific number of segments were chosen randomly from each stratum considering the characteristics of each sampling allocation. We choose the scenario 2 to determine the numbers of segments for each stratum for both allocations (considering an *a priori* estimation provided by the utility manager). Table 6-6 provides the percentage of simulations for each sampling method whose p (the sample estimate of P) is between $P \pm 0.03$.

Table 6-6: % of simulations whose p is between $P\pm0.03$

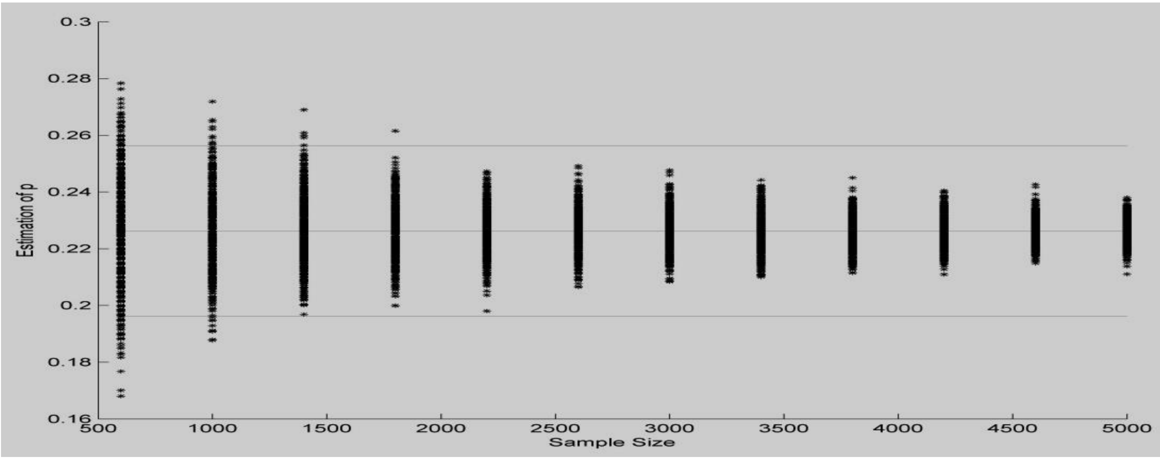
n	Simple random sampling	Proportional allocation in stratified random sampling	Optimum allocation in stratified random sampling
600	92.4	91.7	93.3
1000	98.6	97.9	98.9
1400	99.8	99.5	99.6
1800	99.9	99.9	99.9
2200	100	100	100
2600	100	100	100
3000	100	100	100
3400	100	100	100
3800	100	100	100
4200	100	100	100
4600	100	100	100
5000	100	100	100

The following conclusions could be drawn:

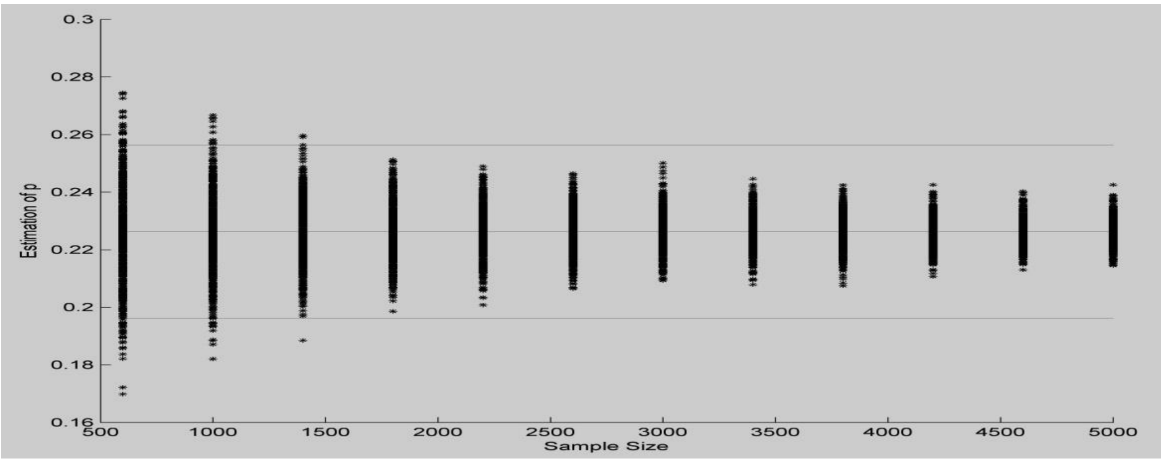
- According to figure 6-3 and as expected, the precision of estimation of P increases with the increase of sample size.
- For $n=1000$, for all sampling methods, the proportion of simulations in which the estimated p is between $P\pm0.03$ is more than 95%.
- In our case, using stratified sampling methods does not improve the estimation compared to the simple random sampling.
- Considering the fact that stratified random sampling does not improve significantly the results, and difficulties to gather certain and complete data on material; the simple random method appears to be the best suited method to design our survey in our case.



Simple random sampling (SRS)



Proportional allocation (PSRS)



Optimum allocation (OSRS)

Figure 6-3: results of 1000 simulations for estimating P, the middle line indicates the *true* value of $P=0.2263;$ two other lines indicates $P\pm0.03$

6.5.2. Impact of calibration sample on the logistic regression

In order to respond to questions 2 & 3 evoked in section 6.2, the results of these simulations were evaluated relative to the model fitted to the population. For converged simulations, the accuracy, precision and statistical significance of each coefficient were also evaluated.

6.5.2.1. Simulation convergence & normality test

Figure 6-5 shows the percentage of simulations which are converged for each sampling method. For simple random sampling (SRS) and proportional allocation in stratified random sampling (PSRS), the percentages of simulations converged are almost equal. However, for optimum allocation in stratified random sampling (OSRS) this percentage is always smaller than the other sampling methods.

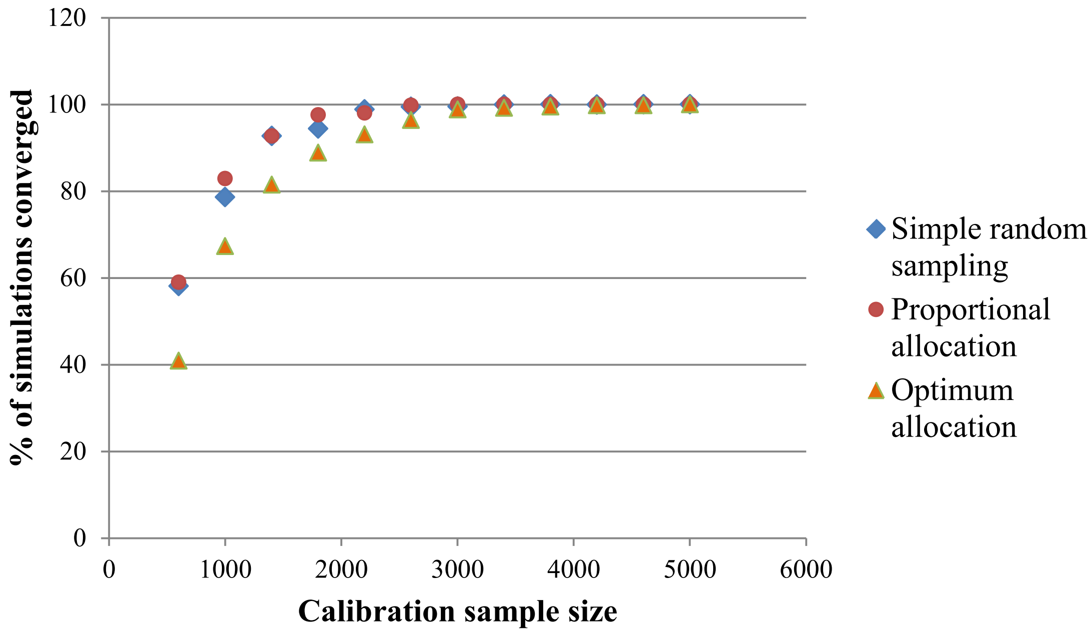


Figure 6-4: Percentage of simulations converged vs. the calibration sample size

Figure 6-5 shows the effect of calibration sample size on the frequency distribution of the values of the regression coefficients for the variable age considering the converged simulations. As the calibration sample size decreases, the distributions become more dispersed and "flatter", particularly when $n < 1000$. In other words, the mode of the distributions becomes flatter not only as the number of converged simulations decreases but also as the domain of the distribution increases (difference between the maximum and the minimum calculated values for a given n). For example, the maximum and minimum values of the regression coefficient were 0.0243 and 0.0186 at $n=3000$ compared with 0.0346 and 0.0117 at $n=600$ for SRS. Therefore, inaccurate estimation of the actual regression coefficient values was more probable at small n. Similar behavior was observed for all other variables.

For small values of n, using SRS or PSRS methods provides better estimates of the *true* value of the population (here equal to 0.0219) than OSRS method as the frequency of distributions' modes increases. For high values of n, all sampling methods provide similar results.

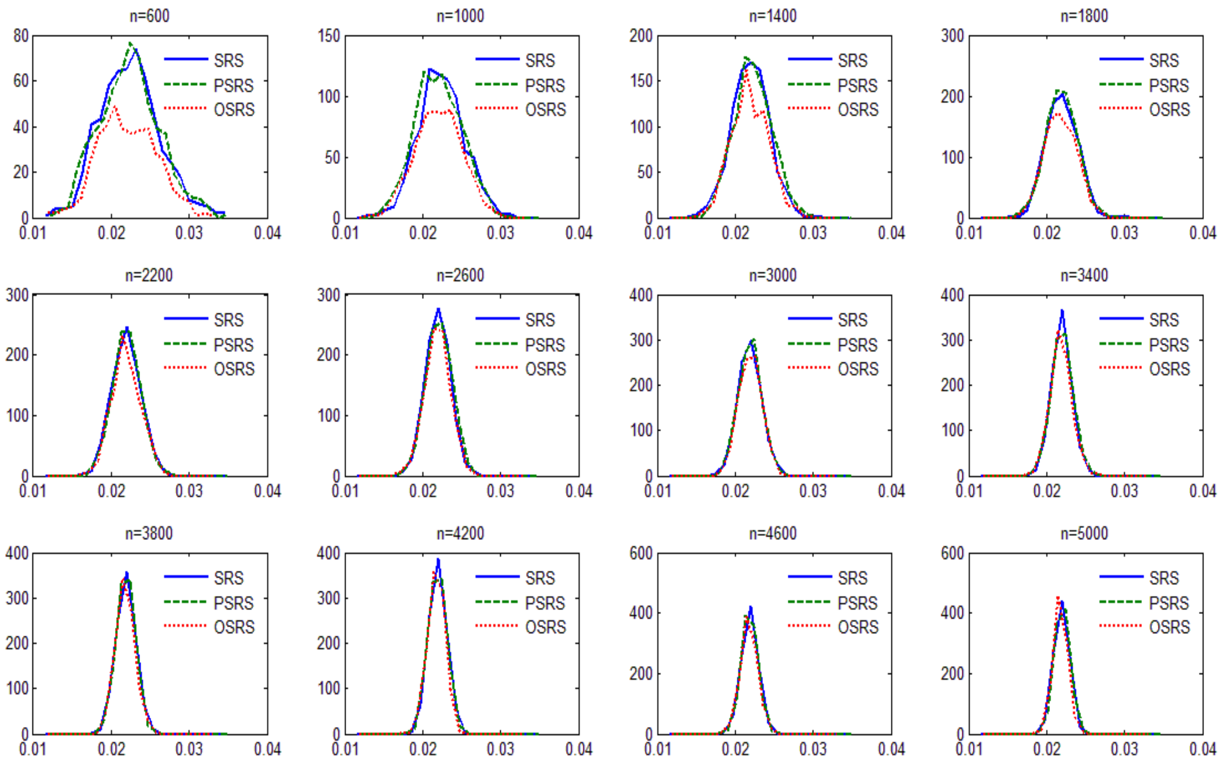


Figure 6-5: Frequency distribution of estimated regression coefficients for variable age according to the calibration sample size. The *true* value for factor age is equal to 0.0219. (Vertical: frequency, Horizontal: coefficient's value)

6.5.2.2. Accuracy of coefficients

Accuracy of coefficients was studied by assessing the average relative bias for each of the regression coefficients considering only the converged simulations (equation 11). Average relative bias decreases with increasing calibration sample size n. Figure 6-6 shows the average relative bias for all variables. Factors depth and material M4 (reinforced concrete) have the highest bias for all *n*. This is due to their relatively high *p*-values.

The average relative bias for materials M1 & M3 (brick and clay tile), by optimum allocation in stratified random sampling is higher than two other sampling methods. This is probably due to the smaller number of segments in brick and in clay tile within the drawn sample by OSRS compared to other sampling methods.

It should be noted that optimum allocation in stratified sampling is designed in order to be optimum about the variable of interest and not for the calibration outcomes of a multi-variable model. Hence, attention should be paid to the fact that this name is chosen within scientific research and does not mean that it is always better than other sampling methods.

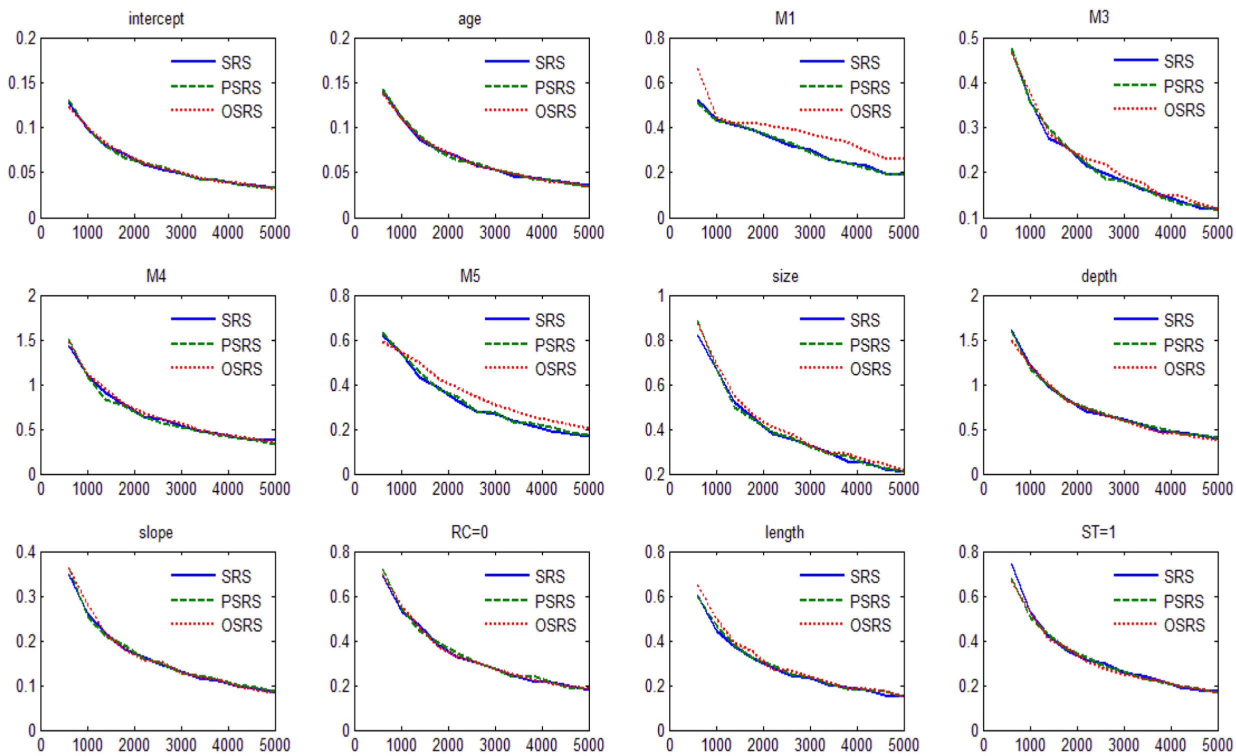


Figure 6-6: average relative bias of estimated regression coefficients (vertical) versus the calibration sample size (horizontal). Where M1: Brick, M2: vitrified clay, M3: clay tile, M4: reinforced concrete, M5: concrete, RC : road class, ST : sewer type (1 : combined).

6.5.2.3. Precision of coefficients

The precision of the estimated regression coefficient was expressed in standard errors (SEs). The SE is calculated for each simulation converged. Afterwards, the ratio of the average SE of coefficient k (over R converged simulations) and the *true* SE obtained from the population (full model) is assessed. This ratio approaches 1 when the SE is estimated correctly. This ratio is calculated by equation 12. Figure 6-7 shows the evolution of this ratio for each sample size and for each sampling method according to the coefficient in question.

The ratios for all coefficients tend to 1 when the sample size increases. However, similar behavior to the average relative bias is observed again for material=brick and clay tile. The corresponding ratios of the optimum allocation in stratified random sampling were higher than two other sampling methods. This is again probably due to the smaller number of segments in brick and in clay tile within the drawn sample by OSRS compared to other sampling methods.

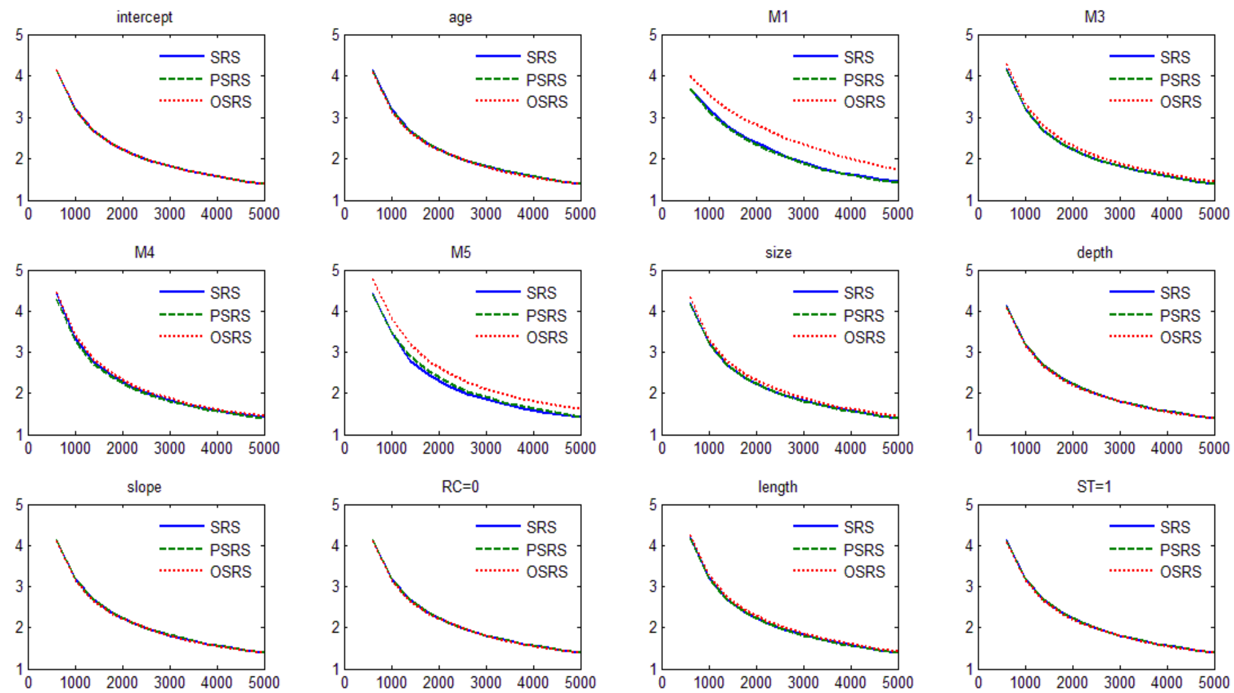


Figure 6-7: the ratio of the average standard errors of simulations divided by the *true* standard error from population (vertical) versus the calibration sample size (horizontal). Where M1: Brick, M2: vitrified clay, M3: clay tile, M4: reinforced concrete, M5: concrete, RC : road class, ST : sewer type (1 : combined).

6.5.2.4. Statistical significance of coefficients

For evaluating the statistical significance of regression coefficients, the proportion of converged simulations in which the 90% confidence interval does not meet zero and contains the *true* value was assessed in the first place. Figure 6-8 shows these proportions for all coefficients.

In general, this proportion increases when the sample size increases (figure 6-8). However, for factor age even with a small sample size, more than 90% of converged simulations produce correct confidence intervals omitting the errors explained in section 6.1. Moreover, this proportion for factors depth and material: reinforce concrete is low as they have important *p*-values. On the other hand, the samples drawn according to SRS or PSRS produce more reliable results in terms of confidence interval compared to OSRS.

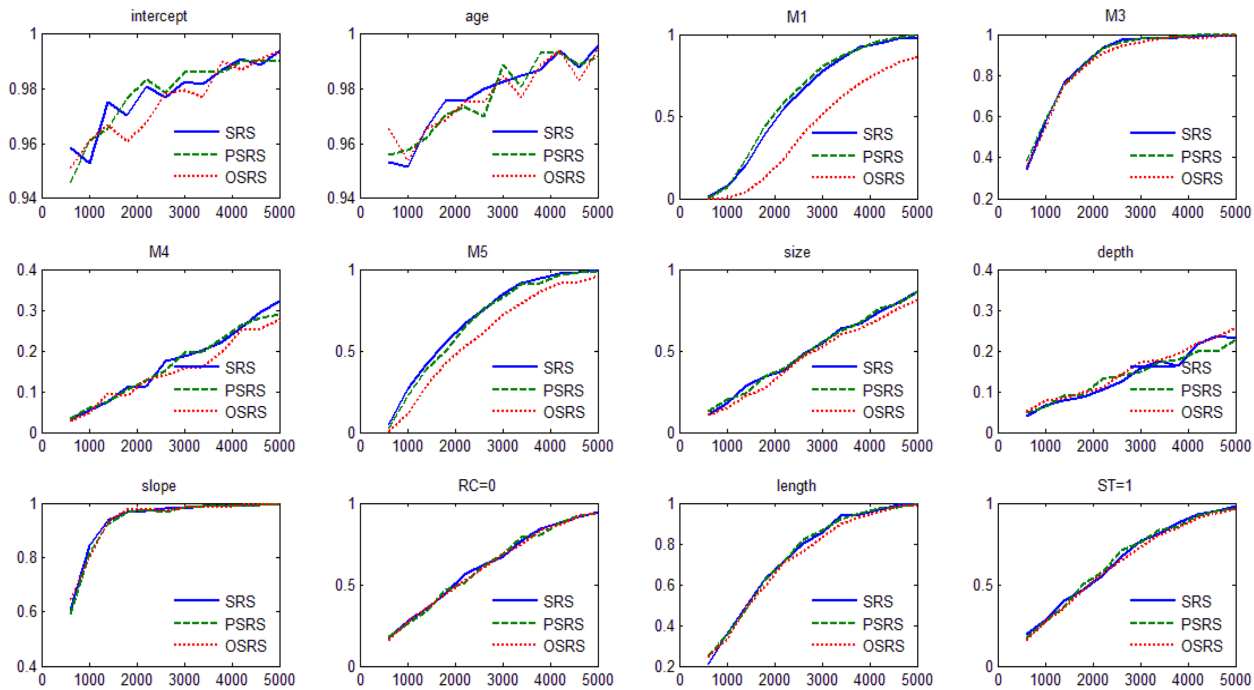


Figure 6-8: proportion of simulations in which the 90% confidence interval about the estimated regression coefficient included the true value and did not meet zero (vertical) versus the calibration sample size (horizontal). Where M1: Brick, M2: vitrified clay, M3: clay tile, M4: reinforced concrete, M5: concrete, RC : road class, ST : sewer type (1 : combined).

Second, the proportion of simulations in which the *p*-value for each coefficient is significantly different from zero (*p*-value <0.1) was evaluated (evaluating type I & III of the possible errors). Table 6-7 provides the number of simulations that converged and the overall percentage of occasions in which each variable was significant at the 10% level under the null hypothesis ($\beta = 0$) respectively for SRS, PSRS and OSRS.

The results show that the percentages of simulations in which coefficients are significantly different from zero for all sampling methods have similar behavior. They increase as the sample size used for calibration of the multivariate model (logistic regression) greatens. However, the results are slightly better for SRS and PSRS compared to OSRS. Table 6-8 provides the results obtained from simulations where PSRS is applied. This table shows that for samples whose sizes are 600, 1000, 1800, 2600, 4200 & 5000, respectively only 1, 2, 3, 4, 8 and 9 factors are significantly different from zero in 90% of cases.

Table 6-7: percentage of occasions in which variable's coefficient was significantly different from *0* amongst converged simulations

variable	*n=600*			*n=1000*			*n=1400*			*n=5000*		
Sampling method	*SRS*	*PSRS*	OSRS	*SRS*	*PSRS*	OSRS	*SRS*	*PSRS*	OSRS	*SRS*	*PSRS*	OSRS
Intercept	100	100	100	100	100	100	100	100	100	100	100	100
Age	100	100	100	100	100	100	100	100	100	100	100	100
M1	6	3	0	23	24	3	44	49	20	100	100	95
M3	52	52	52	73	73	71	87	87	85	100	100	100
M4	10	11	10	13	16	15	18	18	20	50	49	47
M5	20	17	5	48	41	31	60	58	46	100	100	99
Size	23	24	24	34	34	27	43	39	35	94	94	91
Depth	15	15	14	15	16	16	19	18	19	41	37	43
Gradient	75	73	75	92	90	90	98	97	97	100	100	100
RC0	30	29	29	42	40	41	49	50	51	98	98	99
Length	34	39	36	52	51	48	64	62	65	100	100	100
ST	32	32	27	44	42	42	55	54	52	99	99	99
NO of converged simulations	581	590	409	786	829	673	927	927	815	1000	1000	1000

Table 6-9: percentage of occasions in which variable's coefficient was significantly different from *0* amongst converged simulations in PSRS

variable	*n=600*	*n=1000*	n=1400	*n=1800*	*n=2200*	*n=2600*	*n=3000*	*n=3400*	*n=3800*	*n=4200*	*n=4600*	*n=5000*
Intercept	100	100	100	100	100	100	100	100	100	100	100	100
Age	100	100	100	100	100	100	100	100	100	100	100	100
M1	3	24	49	66	77	84	92	95	98	99	100	100
M3	52	73	87	93	98	99	100	100	100	100	100	100
M4	11	16	18	20	25	28	35	34	38	41	46	49
M5	17	41	58	68	79	89	93	97	98	100	100	100
Size	24	34	39	48	57	63	71	76	81	87	90	94
Depth	15	16	18	21	22	25	27	31	28	34	36	37
Gradient	73	90	97	99	100	100	100	100	100	100	100	100
RC0	29	40	50	61	65	74	83	88	89	95	96	98
Length	39	51	62	75	83	90	94	97	98	99	100	100
ST	32	42	54	65	71	82	88	92	93	97	98	99
NO of converged simulations	590	829	927	976	980	998	1000	1000	1000	1000	1000	1000

Figure 6-9 provides the number of simulations in which all coefficients are significantly different from zero versus sample size used for calibrating the logistic regression. It should be noted that even for converged simulations, small number of them could provide same conclusions as the whole population in terms of coefficients significantly different from zero. Therefore, more attention should be given to elimination of some factors from the whole database just because their coefficients were not significant by using a limited number of segments in order to calibrate our multivariate model**.** For example Davis *et al.*, (2001) by calibrating the binary logistic regression on a so-called representative sample of the asset stock of Thames Water, UK, concludes that age, depth, traffic type and road type are not significantly different from zero. They also comment that age is marginally insignificant. However, we showed that perhaps by increasing the sample size, these factors could become significant. Therefore, the sample size is very crucial in terms of model calibration.

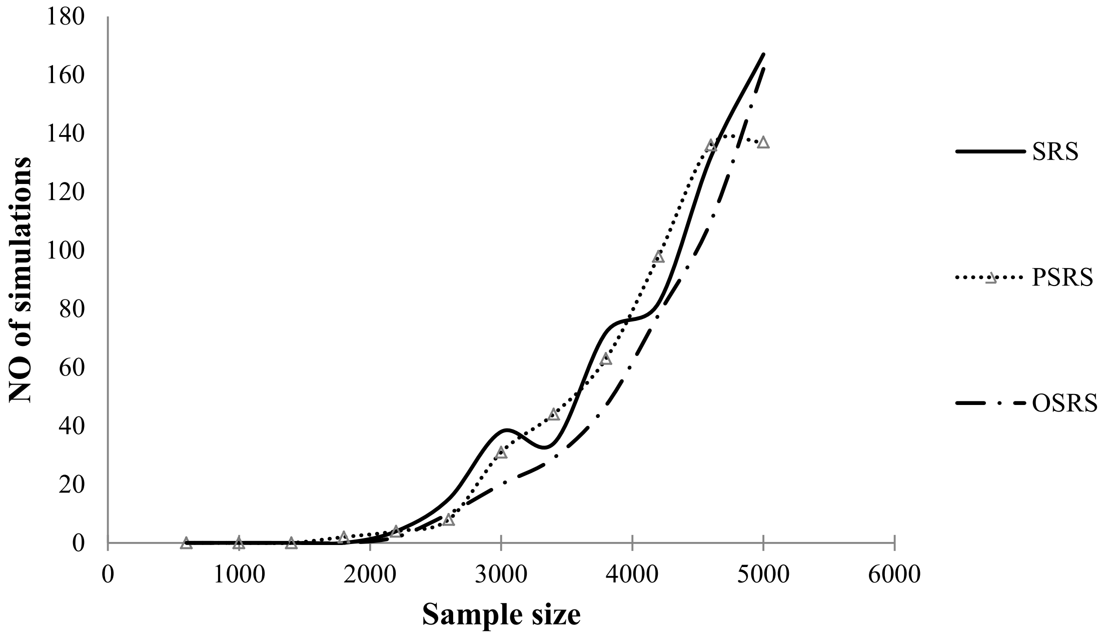


Figure 6-9: Number of simulations in which all coefficients are significantly different from zero

6.6. Conclusions

Within the scientific literature dedicated to the asset management of water and wastewater networks, researchers have showed the importance of using multivariate models in order to use them as tools of decision-making and long term analysis which would result, at the end, to the budg*et al*location plans. However, erroneous conclusions could be drawn about these multivariate models if a proper sample is not used in order to calibrate them. To this end, they have been trying to define and use a representative sample of an asset stock. This notion could be found in articles such as Baur and Herz (2002) and Davies *et al.* (2001).

By definition, a sample which is an appropriate image of an asset stock is the representative sample of this asset stock. In other words, a representative sample is a good sample of a population in the sense that characteristics of interest in the population can be estimated from the sample with a known degree of accuracy (Lohr 2010). However, for instance no study tackles with the problem of representativeness of the used samples in order to calibrate deterioration models in sewer asset management.

The main scopes of this chapter could be summarized in the following questions:

- How to draw a representative sample of an asset stock which reflects in the best and appropriate manners the characteristics of this asset stock?
- How to provide a reliable estimation of a specific property of the asset stock from this sample such as the proportion of segments in failure state (condition grade 5 according to WRc for example)?
- Calibration of multivariate models seems somehow tricky. Therefore, what is the impact of used sample on the calibration outcomes of these multivariate models?

In this chapter, we used three sampling methods not only to calculate the minimum size of a sample in order to be representative of our asset stock but also to draw samples with different sizes (beyond the required size of representativeness) in order to study the influence of size and used sampling method on the calibration process of a multivariate model (we used the logistic regression). We tested simple random sampling and proportional and optimum allocations in stratified random sampling.

The results show that in our case, by considering all sampling methods' requirements, a sample whose size is equal to 1000 segments could be representative of the asset stock. We also showed that giving an estimation of the proportion of segments in failure state by this sample for the whole asset stock is very likely. Moreover, once the sample size increases, this chance greatens as well. According to the results obtained from applied sampling methods and considering the requirements and difficulty of implementing each sampling method, we conclude that in our case the best sampling method is simple random sampling.

However having a randomly-drawn segment whose size is equal to 1000 (assumed representative of the asset stock) does not necessarily mean that the multivariate model's calibration is always successful. Indeed, we showed that even by having a sample of 5000 segments from an asset stock of 9810 segments, the calibration process does not always provide satisfying results for the logistic regression. Therefore, more attention should be given to elimination of some factors from the whole database just because their coefficients were not significant by using a limited number of segments in order to calibrate our multivariate model. However, following points require further investigations:

- In reality, inspections are not carried out randomly and they have a specific motive. Therefore, investigations are required in order to consider the possibility of drawing a representative sample of an asset stock from the available inspections.
- Sampling methods are not limited to those that we have tested. Other sampling methods such as cluster sampling or the combination of them such as two-stage stratification could be considered.
- Some procedures should be proposed in order to identify and to remove problematic variables from the database which could have a serious consequence on the calibration outcomes of the used multivariate model.
- In some cases, because of the nature of available data and information, some multivariate models are not applicable. For example, for small databases, the use of Artificial Neural networks has been criticized for the problem of over-learning (See Tran *et al.*, 2009). Hence, the question here is how to consider the role of available data on the choice of multivariate model.

6.7. References

1. Agresti, A., (2002). “Categorical data analysis”, 2nd Ed. Hoboken, NJ: Wiley.
2. Ahmadi, M., Cherqui, F., Aubin, J.B., De Massiac, J.C., Le Gauffre, P. “Sewer asset management: impact of sample size and its characteristics on the calibration outcomes of a decision-making multivariate model”. Submitted to *Water research Journal*.
3. Allbee, S., (2005). “America's pathway to sustainable water and wastewater systems”. Water Asset Management International, pp. 9-14.
4. ASCE “2013). “Report card for America’s infrastructure.” <http://www.infrastructurereportcard.org/> (Jun. 20. 2013)
5. Bagley SC, White H, Golumb BA, (2001). “Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain”. Journal of Clinical Epidemiologic; 54:979-85.
6. Baur, R. and Herz, R. (2002): “Selective inspection planning with ageing forecast for sewer types”. Journal of Water Science and Technology. 46, 4-7, 389-396.
7. Cochran, William G. (1977). “Sampling Techniques”, 3rd Edition., John Wiley, New York.
8. Concato, J., Feinstein, AR., Holford, TR., (1993). “The risk of determining risk with multivariable models”. Journal of Annals of Internal Medicine; 118: 201-210.
9. Davies, J. P., Clarke, B. A., Whiter, J. T., Cunningham, R. J. and Leidi, A., (2001). “The structural condition of rigid sewer pipes: A statistical investigation”. Urban Water Journal, 3(4), 277–286.
10. Dirksen, J., Clemens, F. H. L. R., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K., Snaterse, C. T. M., (2013). “The consistency of visual sewer inspection data”, Structure and Infrastructure Engineering. Vol. 9, Is. 3, pp. 214-228. (DOI: 10.1080/15732479.2010.541265).
11. EPA (2010): “State of Technology for Rehabilitation of Wastewater Collection Systems”, United States, Environmental Protection Agency, Office of Research and Development, EPA/600/R-10/078, July 2010.
12. Harrell, F.E., Lee, K.L., Matchar, D.B., Reichert, T.A., (1985). “Regression models for prognostic prediction: advantages, problems, and suggested solutions”. Cancer Treatment Reports; 69:1071–7.
13. Le Gauffre, P., Joannis, C., Vasconcelos, E., Breysse, D., Gibello, C. and Desmulliez, J. J. (2007). “Performance indicators and multicriteria decision support for sewer asset management”. Journal of Infrastructure Systems., 13 (2), 105–114.
14. Le Gauffre, P., Joannis, C., Breysse, D., Gibello, C. and Desmulliez, J.J., (2004). "Gestion patrimoniale des réseaux d’assainissement urbains", Guide méthodologique. Paris, Lavoisier Tec&Doc, 416 p., ISBN 27430-0748-6, (in French).
15. Lohr S. L. (2010). “Sampling: Design and Analysis”. Cengage Learning. Boston, USA.
16. Marlow, D., Davis, P., Trans, D., Beale, D., Burn, S., Urquhart, A., (2009). “Remaining asset life: a state of the art review”, WERF – Water Environment Research Foundation, IWA Publishing, London, UK.
17. Marlow, D., Heart, S., Burn, S., Urquhart, A., Gould, S., Anderson, M., Cook, S., Ambrose, M., Madin, B. and Fitzgerald A., (2007). “Condition assessment strategies

and protocols for water and wastewater utility assets", WERF - Water Environment Research Foundation, IWA Publishing, 476 p.

18. Peduzzi, P., Concato, J., Kemper, E., Holford, TR., Feinstein, A.R., (1996). "A simulation study of the number of events per variable in logistic regression analysis". Journal of Clinical Epidemiologic; 49: 1373–1379.
19. Peduzzi, P., Concato, J., Feinstein, A.R., Holford, T.R., (1995). "The importance of events per independent variable (EPV) in proportional hazards regression analysis: II. Accuracy and precision of regression estimates". Journal of Clinical Epidemiologic; 48: 1503-1510.
20. Salman, B., and Salem, O., (2012). "Modeling failure of wastewater collection lines using various section-level regression models", Journal of Infrastructure Systems, ASCE, 18(2), 146-154.
21. Salman, B., (2010). "Infrastructure management and deterioration risk assessment of wastewater collection systems". Thesis (PhD), University of Cincinnati, OH.
22. Sheaffer, R.L., Mendenhall, W.I., Ott, R.L., Gerow. K., (2012). "Survey Sampling", 7th edition. Cengage Learning. Boston, USA.

Chapter VII: Conclusion and perspectives

An important phase of sewer asset management is the evaluation of physical condition of assets. This latter relies mostly on visual inspections. Amongst visual inspection methods, the closed circuit televised inspections are mainly used and applied to the sewer segments. While it is widely used, the assessment of a condition grade to a sewer segment is being carried out by applying and adapting different condition grading protocols all over the world.

Although most protocols give a segment a condition grade by comparing its score with a subjective scale of numerical values, we propose a protocol to calibrate thresholds for each asset stock considering intentions of utility managers about the under or overestimation of assets' condition grades.

However, the condition grading of segments should be done after inspecting them. Nevertheless, high number of deficient segments resulting from poor past maintenance activities and budget restrictions necessitate the development of prioritization tools to address inspection needs of the segments with the highest risk of failure.

Accordingly, the efficiency of sewer inspection programs improves if inspection survey targets more problematic segments. Hence, it seems that accurate predictions of the current and future condition of sewers are crucial. These predictions can be obtained from deterioration models. Though various types of deterioration models may be found, attention is still focused on the type of deterioration model used and the influence of available data on the predictive power of these models has gone unstudied.

On the other hand, at the current situation the use of a representative sample from an asset stock in order to calibrate decision-support models as deterioration models, to study scenarios about future and to draw appropriate conclusions for the whole asset stock seems mandatory as small number of utilities has inspected the totality of their assets.

The objectives of this thesis, therefore, were as follows:

- To propose a framework for inspection programs and to study the influence of the quality and availability of data on these programs;
- To test the sensitivity of condition grading protocol proposed and developed during the French projects RERAU and INDIGAU to the parameters needed to be fixed by utility managers;
- To draw a representative sample of an asset stock which reflects in the best and appropriate manners the characteristics of this asset stock;
- To study the impact of used sample on the calibration outcomes of these multivariate models.

7.1. Inspection programming and influence of data on its efficiency

In this manuscript, we proposed a systematic approach using a deterioration model to improve sewer inspection programs efficiency. The logistic regression was selected in order to predict the condition grade of sewer segments by considering their influential factors.

Afterwards, by considering a semi-virtual asset stock, we studied the influence of incompleteness, imprecision and uncertainty within data on the efficiency of inspection programs. We showed that the quality of data in terms of imprecision, incompleteness and uncertainty could have major consequences on the efficiency of inspection programs. It is found that imprecise or uncertain data are better than not having them at all. However, if data do not exist (are not available), some auxiliary variables could be considered in order to compensate effects of their absence on the inspection programs.

We also elaborated the list of most informative single factors amongst available influential factors by proposing 2 different approaches:

- A backward selection method using the deviance statistic (likelihood ratio test) and adapting a nested subset process;
- A forward selection method applying the developed systematic approach for inspection programs based on an indicator to gauge the efficiency of inspection programs.

The results for two approaches were similar. This makes it possible to use the first approach based on the deviance statistic on a representative sample of an asset stock to establish the list of the most informative single factors and also to plan the data acquisition programs.

As perspectives, following points could be considered even though the first one was partially treated within chapter 6:

- The influence of initial stock of segments which is used in order to calibrate the deterioration model;
- We showed that the backward selection method based on the deviance statistic produces similar results as forward selection method based on inspection programs framework for the whole asset stock. In other words, this approach was applied on the whole asset stock. However, in the reality, this is not the case. Therefore, we should study outcomes of the application of this method on a representative sample of our asset stock and compare the results to see if similar results are again produced.
- An economic analysis of inspection programs defined in order to gauge benefits that a utility could make by defining these programs;

- There is no solid recommendation concerning the choice of deterioration model within the scientific literature. Hence, this choice could be the objective of further investigations.

7.2. Assessing a segment by an inspection

Once a segment is inspected, it should be evaluated. Although visual inspection is widely used in order to assess assets' conditions, many sources of uncertainty exist within the process of elaborating rehabilitation programs from the results of investigations.

We have identified 4 main uncertainty sources within the evaluation process:

- Inspecting a segment and saving the existing defects by codes defined by sewer inspection norms such as PACP, EN 13508-2 or WRc by CCTV operator.
- The conversion process of codes into a score for each segment. This conversion depends on the severity, shape and length of defects which could be taken into account by a peak score, total score or mean score calculated for a given segment.
- Assigning a condition grade to a segment from its score considering some thresholds. For example for two segments evaluated by WRc protocol, if the corresponding scores are successively 80 and 164, both of them are assigned into G4. However, the nature and severity of defects could be very different. Therefore, an expert may classify the latter into a more sever condition grade.
- The fourth type of uncertainty is related to the table used to aggregate an indicator derived from visual inspection with another indicator such as risk factor or vulnerability of the environment (for assessing a rehabilitation criterion). Crossing indicator X (grade G2) with indicator Y (grade G3) to assess indicator Z may lead to two possible crisp aggregation results: G2 by an optimistic and G3 by a pessimistic view. Which option should be chosen?

In this thesis, we tackled the 3rd source. Though there are many condition grading protocols used and applied all over the world, almost all of them fail to tackle the problem of uncertainty within the evaluation process. Hence, the condition assessment process for a segment could result in an over-estimation or an under-estimation of the real condition grade of the segment. Accordingly, used condition grading protocols do not allow the stakeholders to take into account their own specifications. Therefore, the need of a more specific condition grading protocol is felt which could take into account the sensitivity of utility managers and stakeholders to this specific issue.

Consequently, a condition grading protocol and procedure was proposed within French projects RERAU and INDIGAU based on a sample of chosen segments from many French utilities evaluated by experts for defined dysfunction indicators within these projects such as infiltration, collapse, chemical attack and etc.

Additionally, these projects propose, for each dysfunction indicator, a quantification table allowing the translation of observed defects into a mean score. Once all these segments, evaluated by experts from different French utilities, are assessed by the proposed quantification table, a cost criterion fixes the required thresholds in order to assign a condition grade for each dysfunction indicator to each segment. This criterion is based on two parameters:

- The overall condition grade of the asset stock in-question;
- The assignment-error weights which would represent the sensitivity of utility managers to the over or underestimation of assets' condition grade.

In this thesis, we also carried out some sensitivity analyses of these parameters. We showed that both parameters could influence the thresholds and accordingly the proportion of segments in each condition grade. This latter is shown by applying found thresholds on 4471 segments of the Greater Lyon asset stock.

However, further investigations should be performed in order to deal with following issues:

- Improvement of the calibration requirements:
  - Increasing the size of the sample evaluated by experts. For instance, 60 segments are evaluated by experts for each defined dysfunction indicator.
  - Replacing discrete assessments of experts (figure 7-1) by a continuous distribution for each condition grade
- Identification of defect codes which have most important influence on the condition assessment procedure by taking into account the quantification tables used for each dysfunction indicator;
- A more in-depth sensitivity analysis of used parameters. As we have discrete evaluations of experts (figure 7-1 for indicator *Infiltration)*, thresholds may vary significantly (jump between discrete evaluations) just by a small modification in parameters.

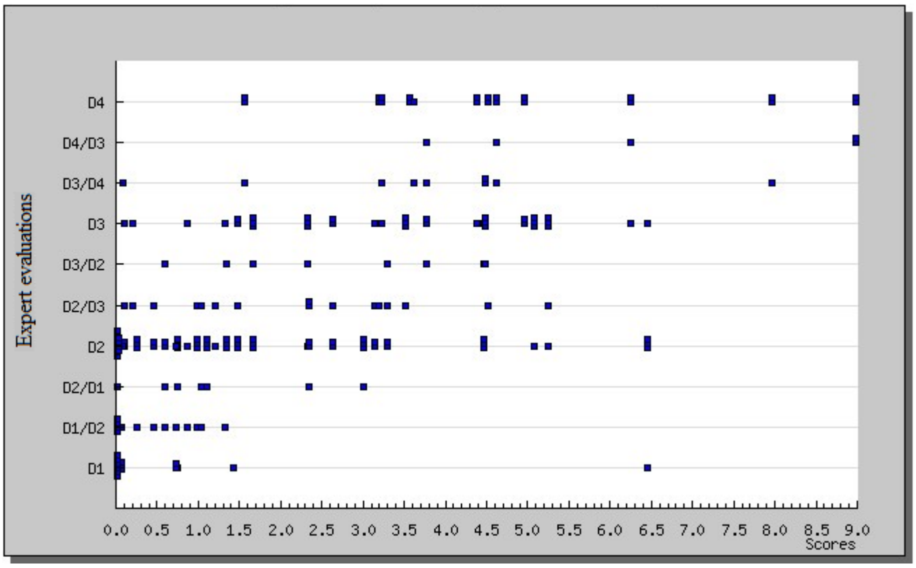


Figure 7-1. Distribution of expert evaluations vs. calculated scores (*a*=3 and *b*=1)

7.3. Assessing an asset stock from a sample

"Assessing an asset stock from a sample", this expression covers several possible management objectives. The first possible objective (a) is to estimate the current condition of an asset stock and to estimate current rehabilitation needs. The second possible objective is to calibrate failure models that may be used (b) in the definition of inspection priorities and (c) in the estimation of future rehabilitation needs due to aging.

In fact, at the moment small number of utilities has completely inspected their asset stocks. Therefore, the use of a sample from an asset stock in order to calibrate deterioration models, to study scenarios about future and to draw appropriate conclusions seems mandatory and has been stressed by many researchers. This sample should reflect, however, the characteristics of the asset stock in-question in the best manner. By definition, a sample which is an appropriate image of an asset stock is the representative sample of this asset stock.

On the other hand, within the scientific literature dedicated to the asset management, researchers developed and calibrated deterioration models without paying attention to the impact of used sample on the outcomes. The main interest of doing so was to test and compare the deterioration models or to develop a risk-based rehabilitation approach.

Accordingly, following issues should be solved concerning the appropriate use of a representative sample of an asset stock in a decision-making process:

- How to draw a representative sample of an asset stock which reflects in the best and appropriate manners the characteristics of this asset stock?
- How to provide a reliable estimation of a specific property of the asset stock from this sample?
- What is the impact of used sample on the calibration outcomes of these multivariate models?

In order to answer these questions, we tested three sampling methods not only to calculate the minimum size of a sample in order to be representative of our asset stock but also to draw samples with different sizes in order to study the influence of size and used sampling method on the calibration process of a multivariate model (we used the logistic regression). These methods consist of simple random sampling, proportional and optimum allocations in stratified random sampling.

The results show that there is a small difference between sampling methods for drawing a representative sample of the asset stock in terms of required sample size. However, having a representative sample of the asset stock does not necessarily mean that the multivariate model's calibration is always successful. A successful calibration depends, however, on the following points:

- Type of the multivariate model used;
- Sampling method;
- Number of segments inspected;
- Number of segments having characteristic of interest within these inspected segments (for example number of segments in failure state in case of a binary condition grade);
- Number of influential factors available (number of variables);
- Quality of available data;

However, following points require further investigations:

- In reality, inspections are not carried out randomly and they have a specific motive. Therefore, investigations are required in order to consider the possibility of drawing a representative sample of an asset stock from the available inspections.
- Sampling methods are not limited to those that we have tested. Other sampling methods such as cluster sampling or the combination of them such as two-stage stratification could be considered.

- Some procedures should be proposed in order to identify and to remove problematic variables from the database which could have a serious consequence on the calibration outcome of the used multivariate model.
- In some cases, because of the nature of available data and information, some multivariate models are not applicable. For example, for small databases, the use of Artificial Neural networks has been criticized for the problem of over-learning. Hence, the question here is how to consider the role of available data on the choice of multivariate model?

FOLIO ADMINISTRATIF

THESE SOUTENUE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON

NOM : AHMADI DATE de SOUTENANCE : 11 avril 2014
Prénoms : Mehdi

TITRE : Gestion patrimoniale des réseaux d'assainissement : impact de la qualité des données et du paramétrage du modèle utilisé sur l'évaluation de l'état des tronçons et des patrimoines.

NATURE : Doctorat Numéro d'ordre :

Ecole doctorale : Mécanique, Energétique, Génie civil, Acoustique (MEGA)
Spécialité : Génie Civil Environnemental

RESUME :
La gestion patrimoniale est une problématique d'importance croissante pour les gestionnaires des réseaux d'eau potable et d'assainissement. Elle vise à choisir les meilleures décisions d'actions à mener sur les éléments du patrimoine, pour limiter les risques, optimiser les performances et réduire les couts. Cela implique une démarche proactive incluant le développement d'outils de hiérarchisation pour sélectionner les conduites à inspecter et/ou réhabiliter. Dans ce manuscrit, nous avons adapté la définition spécifique de la gestion proactive des réseaux d'assainissement aux pratiques des gestionnaires pour passer de la maintenance post-défaillance à une politique de gestion patrimoniale proactive sur le long-terme. Pour cela, nous avons identifié les verrous suivants qui sont abordés dans le manuscrit.
Premièrement, les inspections des conduites sont actuellement programmées soit selon le jugement du gestionnaire ciblant les secteurs problématiques (odeurs, infiltrations, etc.), soit indépendamment des choix du gestionnaire (lors des travaux de voirie programmés, etc.). Il est donc nécessaire de pouvoir élaborer ces programmes d'inspections à partir de modèles de détérioration pour optimiser les budgets alloués. Malgré le développement de nombreux modèles de détérioration, l'influence de la qualité et de la disponibilité des données (incomplétude, imprécision, incertitude) sur la puissance prédictive des modèles n'a pas été étudiée en détail. Nous avons abordé cette question dans la première partie de la thèse, en proposant deux méthodes pour déterminer la liste des facteurs les plus informatifs à partir d'un échantillon représentatif. Entre autres, nos résultats suggèrent que l'imprécision sur une donnée est préférable à ne pas disposer de cette donnée. Nous avons également montré que la notion de « quartier » pourrait être utilisée, sous certaines conditions, pour compenser la non-connaissance de l'âge des conduites.
Deuxièmement, une fois les conduites inspectées, leur état de santé doit être évalué à l'aide d'un protocole identique pour l'ensemble du patrimoine. Bien que différents protocoles existent, aucun ne permet de prendre en compte les considérations du gestionnaire concernant la sur- ou sous-estimation de l'état de santé (liée aux nombreuses incertitudes induites par l'ensemble du processus). Il est donc nécessaire que le protocole utilisé puisse minimiser les erreurs de sur- ou sous-estimation en fonction des choix des gestionnaires. Nous avons proposé une procédure prenant en compte ces choix, ainsi que l'état global du patrimoine considéré. Les études de sensibilité réalisées à partir des données d'inspection d'une partie du patrimoine du Grand Lyon montrent que l'état de santé évalué dépend fortement des choix faits par le gestionnaire concernant la sur- ou sous-estimation.
Troisièmement, peu de patrimoines ont actuellement été complètement inspectés et évalués. Ainsi, l'utilisation d'un échantillon représentatif d'un patrimoine semble obligatoire pour pouvoir calibrer des modèles justifiant les décisions comme par exemple un modèle de détérioration. Néanmoins, cela pose les problèmes suivants : 1) comment générer un échantillon représentatif reflétant au mieux les caractéristiques du patrimoine complet ? 2) Quel est l'impact de cet échantillon sur les résultats de calage des modèles multi-variables ? Ainsi, après avoir généré différents échantillons de différentes tailles et selon différentes méthodes d'échantillonnage, à l'aide de la méthode de Monte Carlo, nous avons pu proposer une procédure pour étudier l'influence des caractéristiques de l'échantillon sur les résultats du modèle de détérioration. A partir d'analyses statistiques, nous avons montré que le processus de calage (des modèles) dépend fortement de l'échantillon disponible, et cela peut donc conduire à des résultats erronés.

MOTS-CLES : Gestion patrimoniale, assainissement, aide à la décision, programme d'inspection des conduites, état de santé, régression logistique.

Laboratoire (s) de recherche : Laboratoire de Génie Civil et d'Ingénierie Environnementale (LGCIE)

Directeurs de thèse: Pascal Le Gauffre, Frédéric Cherqui

Président de jury :

Composition du jury :

Mme. BARRAUD Sylvie	Professeure (INSA de Lyon)	Examinatrice
M. BREUL Pierre	Professeur (Univ. Blaise Pascal, Clermont-Ferrand)	Examinateur
M. BREYSSE Denys	Professeur (Université Bordeaux 1)	Examinateur
M. CHERQUI Frédéric	Maître de Conférences (Université Lyon 1)	Co-Directeur de thèse
M. CLEMENS Francois	Professeur (TU DELFT)	Rapporteur
M. LE GAUFFRE Pascal	Maître de Conférences, HDR (INSA de Lyon)	Directeur de thèse
M. MAILHOT Alain	Professeur (INRS, Canada)	Rapporteur